Universidad Autónoma de Madrid
Escuela politécnica superior

Doble Grado en Ingeniería Informática y Matemáticas

Trabajo Fin de Grado

HERRAMIENTA INFORMÁTICA PARA EL CÁLCULO DE ÍNDICES DE PODER EN SISTEMAS DE VOTACIÓN PONDERADOS

Autor: Alejandro del Moral Charneco
Tutora: Irene Rodríguez Luján
Ponente: Gonzalo Martínez Muñoz

Julio 2017

HERRAMIENTA INFORMÁTICA PARA EL CÁLCULO DE ÍNDICES DE PODER EN SISTEMAS DE VOTACIÓN PONDERADOS

Autor: Alejandro del Moral Charneco
Tutora: Irene Rodríguez Luján
Ponente: Gonzalo Martínez Muñoz

Dpto. de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Julio 2017

Agradecimientos

En primer lugar, me gustaría agradecer a mi tutora, Irene Rodríguez Luján, todo el esfuerzo y tiempo dedicados. Sin su ayuda no habría sido posible el desarrollo de este trabajo.

También, quiero agradecer a los profesores que me han dado clase a lo largo de estos años los conocimientos que me han transmitido.

Por supuesto, doy las gracias a mis compañeros de carrera y amigos, pues de no ser por ellos no habría llegado tan lejos en este largo camino. Han sido cinco años de duro trabajo, pero con su apoyo han hecho que cada día fuera especial y han conseguido que no me rinda en ningún momento.

Quiero dar las gracias a mi madre, por aguantar mis agobios y por prestarme su ayuda, sobre todo en los últimos días de trabajo. Son muchas horas las que se ha pasado sentada a mi lado, dándome su apoyo y animándome a seguir adelante. También agradezco al resto de mi familia todos sus mensajes de ánimo, que me han dado energía para continuar.

Por último, le agradezco a mi novia que siempre haya estado ahí, apoyándome incluso en los peores momentos. Ha aguantado mi carácter y ha conseguido tranquilizarme siempre que no veía la luz al final del túnel en estos cinco años.

Resumen

Los índices de poder son una métrica que permiten estudiar el poder de decisión de un partido político en una votación. En este trabajo se ha realizado un estudio de tres índices de poder: índice de poder de Borda, índice de poder de Banzhaf e índice de poder de Shapley-Shubik.

Dado que estos realizan agrupaciones de partidos políticos, se ha realizado un estudio de datos de Twitter para obtener probabilidades de agrupación mediante una aplicación en Python.

Teniendo en cuenta que el cálculo de estos índices de poder se basa en realizar permutaciones y combinaciones, se han desarrollado teorías para reducir el número de ellas; desarrollando una teoría propia que permite combinar juegos (3,2) como juegos (2,2), así como teorías de artículos que incluyen abstención en los sistemas de votación sin que esto suponga un incremento descomunal del coste computacional.

Este cálculo se ha realizado mediante una aplicación en R usando las librerías de Shiny y Shinydashboard, pues mejora la usabilidad y experiencia de usuario. Esta aplicación contiene los datos de las elecciones al Congreso de los Diputados desde 1977 hasta 2016. Estos datos han sido introducidos en el propio código de la aplicación, pues la estructura de los archivos no era la más idónea para un buen tratamiento de datos.

Para cada uno de los índices de poder se ha desarrollado un algoritmo que permite calcularlos. En el caso del índice de poder de Banzhaf se han incluido votaciones con mayoría simple y absoluta, permitiendo la abstención. En cambio, en el índice de poder de Borda y el índice de poder de Shapley-Shubik sólo se han permitido votaciones con mayoría absoluta y sin abstenciones.

Tras el análisis de ambas ideas, índices de poder y probabilidades de agrupación, se ha concluido que el índice de Banzhaf y de Shapley-Shubik reflejan el panorama actual de la política española, pues muestran que los partidos políticos con mayor poder son los mismos así como la imposibilidad de sacar probabilidades de agrupación a partir de los datos de Twitter.

Palabras Clave

Sistema de votación ponderado, coaliciones, votante basculante, votante pivotante, cambios a corto, cambios a largo, mayoría simple, mayoría absoluta, juegos (3,2), juegos (2,2), índice de poder de Borda, índice de poder de Banzhaf, índice de poder de Shapley-Shubik.

Abstract

The power indexes are a metric that allow to study the power of decision of a political party in a polling.In this elaboration a study of three power indexes have been carried out: Borda power index, Banzhaf power index and Shapley-Shubik power index.

Due to the fact that these algorithms make groups, a Twitter study based on its data has been conducted to obtain grouping probabilities using a Python application.

Taking into account that the calculation of these power indices is based on making permutations and combinations, theories have been developed to reduce the number of them; developing (3.2) as games (2,2), as well as theories of articles that include abstention in the voting systems without supposing a huge increase of the computational cost.

This calculation has been done by an application in R using the Shiny and Shinydashboard libraries, as it improves usability and user experience. This application contains the data of the elections to the Congress of Deputies from 1977 to 2016. These data have been entered in the application code itself because the file structure was not the most suitable for a good data processing.

For each power index an algorithm which allows calculating it has been developed. In the case of the Banzhaf power index, a simple and absolute majority has been included, allowing abstention. On the other hand, in Borda power index and Shapley-Shubik power index, only absolute majority and no abstention votes were allowed.

After analyzing both ideas, power indexes and grouping probabilities, it has been concluded that the Banzhaf and Shapley-Shubik index reflect the current panorama of Spanish politics, as well as the impossibility of getting grouping probabilities from the Twitter data.

Key words

Weighted voting system, coalitions, swing voter, pivoting voter, short changes, long changes, simple majority, absolute majority, games (3,2), games (2,2), Borda power index, Banzhaf power index and Shapley-Shubik power index.

Índice general

Índice de Figuras XI

Índice de Tablas XII

1. Introducción 1

1.1. Motivación del proyecto . . . 1
1.2. Objetivos y enfoque . . . 2
1.3. Estructura de la memoria . . . 3

2. Marco Teórico 5

2.1. Análisis de redes sociales . . . 5
2.1.1. Twitter . . . 6
2.1.2. Probabilidades de coalición . . . 6
2.2. Índices de poder . . . 7
2.2.1. Sistemas de votación ponderados . . . 7
2.2.2. Índice de poder de Borda . . . 9
2.2.3. Índice de poder de Banzhaf . . . 9
2.2.4. Índice de poder de Shapley-Shubik . . . 11
2.2.5. Comparación entre IB y ISS . . . 14
2.3. Teoría de Juegos (3,2) . . . 14
2.3.1. Juegos sin abstención . . . 15
2.3.2. Juegos con abstención . . . 16

3. Diseño y desarrollo 19

3.1. Herramientas y tecnologías . . . 19
3.1.1. Herramienta en Shiny. Herramienta en Python . . . 19
3.2. El conjunto de datos . . . 21
3.3. Algoritmos . . . 23
3.3.1. Algoritmo desarrollado para el cálculo de índices de poder en juegos (3,2) 23
3.3.2. Algoritmo de Borda . . . 25
3.3.3. Algoritmo de Banzhaf . . . 25
3.3.4. Algoritmo de Shapley-Shubik . . . 28

4. Análisis resultados **31**
4.1. Probabilidades de coalición 31
4.2. Índices de poder 33
4.2.1. Índice de Borda 33
4.2.2. Índice de Banzhaf 34
4.2.3. Índice de Shapley-Shubik 36

5. Conclusiones y trabajo futuro **37**
5.1. Conclusiones 37
5.1.1. Probabilidades de coalición 37
5.1.2. Herramienta de cálculo de índices de poder 37
5.1.3. Propuesta de algoritmos para el cálculo eficiente de índices de poder en juegos (3,2) 37
5.1.4. Análisis y estudio de índices de poder 38
5.2. Trabajo futuro 38

Glosario de acrónimos **39**

Bibliografía **40**

Índice de Figuras

3.1. Interfaz gráfica de la aplicación de cálculo de los índices de poder. 20
3.2. Tabla de la encuesta preelectoral de 2016 del CIS [1] 21
3.3. Ejemplo de fichero de datos de las elecciones al Congreso obtenido de [2] 22
3.4. Esquema del cálculo del índice de Borda para un partido 25

4.1. Histograma del índice de Borda 33
4.2. Histograma del índice de Banzhaf sin abstención 34
4.3. Histograma del índice de Banzhaf con abstención 35
4.4. Histograma del índice de Shapley-Shubik 36

Índice de Tablas

2.1. Combinaciones de votos del ejemplo 2.2 10

3.1. Combinación de participantes x_i 26
3.2. Tabla del ejemplo 3.2 27
3.3. Tabla de permutaciones del ejemplo 3.3 28
3.4. Tabla de valores del índice de poder de Shapley-Shubik 29

4.1. Total de usuarios de Twitter. 31
4.2. Datos de cuentas que siguen a políticos. 31
4.3. Datos de cuentas que siguen a partidos políticos. 32
4.4. Matriz de probabilidades de partidos condicionadas a partidos. 32
4.5. Matriz de probabilidades de políticos condicionadas a partidos. 32
4.6. Matriz de probabilidades de partidos condicionadas a políticos. 32

1

Introducción

1.1. Motivación del proyecto

La política es un tema que está muy de moda actualmente, pero realmente, ¿cómo podemos saber el verdadero poder que tiene cada partido político? Tener un escaño en el congreso no es sinónimo de tener poder, pues no implica tener dominio sobre las decisiones. En realidad, el poder de los escaños se mide en base a la capacidad de los mismos de hacer que una ley se apruebe o no.

Dentro de este contexto, en este Trabajo de Fin de Grado se investigan tres índices de poder distintos comúnmente utilizados para medir el poder de decisión de los jugadores en sistemas de votación ponderados. En particular, este trabajo se centra en medir el poder de los partidos políticos en función de los escaños obtenidos, dado que las elecciones no son más que un juego simple donde los partidos políticos son los participantes, y las opciones en ese juego son si/no/abstención, y podemos tener dos configuraciones del juego, mayoría simple o mayoría absoluta.

Aunque los artículos presentando los índices de poder más comunes datan de los años 1999, este es un campo de investigación aún activo y en el que se siguen haciendo contribuciones hoy en día. El conocimiento del resultado de los análisis de estos índices de poder podría ser de utilidad para los partidos políticos que tendrían la posibilidad de cuantificar el poder de las coaliciones, tanto para detectar aquellos partidos clave en las decisiones como aquellos cuyo voto no supone ningún cambio en la decisión final.

Otro campo de conocimiento altamente relacionado con este TFG y de gran actualidad es el campo de la ciencia de datos cuyo el objetivo último es dar valor a la gran cantidad de información que se recoge hoy en día. Más concretamente, en este TFG se han recogido datos de la plataforma Twitter al ser una red social que los partidos políticos utilizan para hacer campaña electoral. Se ha realizado un estudio de los datos de Twitter para analizar la viabilidad de estimar las probabilidades de coalición entre partidos políticos utilizando la información procedente de los *followers*.

1.2. Objetivos y enfoque

El objetivo principal de este Trabajo de Fin de Grado es llevar a cabo un estudio de las teorías y métricas más utilizadas en la literatura para la cuantificación de índices de poder de partidos políticos así como desarrollar una aplicación para su cálculo.

En un principio, la aplicación estuvo pensada para calcular los índices de poder de los partidos políticos en función a los escaños obtenidos en las elecciones al Congreso de los Diputados, obteniendo los datos de la página web del Ministerio del Interior [2]. Algunos ejemplos de votaciones ponderadas son las votaciones del Congreso de los Diputados o las votaciones del Senado, donde el sistema de votación consiste en participantes con un voto cada uno, que puede tener un valor de sí/no/abstención. De esta forma, para que una ley sea aprobada se debe superar la cuota, que son los votos mínimos para aprobar una ley. Por simplicidad, se supondrá que cada diputado, senador o participante de un mismo partido político o agrupación votará lo mismo que sus compañeros, es decir, o todos sí, no o abstención. Posteriormente se decidió ampliar la funcionalidad de la aplicación para poder calcular el índice de poder de cualquier participante en una votación ponderada no necesariamente dentro del ámbito político. Por ejemplo, en un juego online calcular el poder de cada participante en la coalición utilizando el nivel, poder de ataque, o cualquier dato equivalente a los escaños en los partidos políticos.

Además de llevar a cabo el cálculo de los índices de poder en base al número de votos de cada participante, se planteó como objetivo de este TFG realizar un estudio con datos obtenidos a partir de Twitter para calcular probabilidades de coalición entre partidos políticos, dado que las teorías estudiadas asumen coaliciones equiprobables, situación que parece poco realista en el caso de la política española al haber coaliciones que podrían considerarse como prácticamente imposibles. Para realizar este estudio se implementó un programa en Python que semanalmente almacena los followers de ciertas cuentas de Twitter que se han considerado relevantes y en base a esta información se estimaron probabilidades de coalición entre partidos. En el apartado 2.1.2 se analizará con más detalle la metodología y en la sección 5.1.1 los resultados obtenidos en este sentido.

Por tanto, los objetivos de este Trabajo de Fin de Grado se pueden desglosar en los siguientes puntos:

1. Estudio y análisis teórico de cada uno de los tres índices que se han considerado más importantes. Estos son el índice de Borda, de Banzhaf y de Shapley-Shubik.

2. Análisis de los datos de Twitter para estimar cuantitativamente las probabilidades de coalición entre partidos políticos.

3. Diseño e implementación de una aplicación web que calculará los tres índices de poder y mostrará gráficamente los resultados obtenidos de acuerdo a diferentes métricas. Esta aplicación está pensada para trabajar tanto con datos importados correspondientes a las elecciones del Congreso de los Diputados desde 1977 hasta 2016, como con datos que el usuario puede introducir manualmente.

4. Análisis de los resultados obtenidos a partir de diferentes métricas generadas por la aplicación web. Estas métricas y análisis se centran principalmente en la cuantificación del índice de poder de cada participante de una votación ponderada con configuración de mayoría absoluta como mayoría simple. También se estudiará la diferencia de poder que implica pasar de un juego de mayoría absoluta a un juego de mayoría simple sin modificar los datos de entrada.

1.3. Estructura de la memoria

En base a los objetivos descritos en la sección anterior, esta memoria se divide en los siguientes capítulos:

- **Capítulo 1:** Introducción
- **Capítulo 2:** Marco teórico. En este capítulo se presentarán los índices de poder considerados en este trabajo así como la obtención de probabilidades de coalición basados en datos de Twitter. También se introduce la teoría de juegos (3,2) que permite introducir sistemas de votación con abstención.
- **Capítulo 3:** Diseño y desarrollo. En este capítulo se describen las herramientas y tecnologías desarrolladas, el conjunto de datos utilizados para este trabajo así como los algoritmos desarrollados para el cálculo de los índices de poder.
- **Capítulo 4:** Análisis de resultados. En este capítulo se comentarán los resultados obtenidos mediante gráficas correspondientes al poder de cada partido así como las matrices de probabilidades de coalición.
- **Capítulo 5:** Conclusiones. Por último, en este capítulo se discuten las conclusiones derivadas de la realización del Trabajo de Fin de Grado y se presentan posibles líneas de trabajo futuro.
- **Glosario, Bibliografía y Anexos**

2
Marco Teórico

En este trabajo vamos a estudiar, implementar y analizar los índices de poder de participantes en sistemas de votación ponderados (SVP), es decir, un entorno de toma de decisiones en el que los participantes tienen un número diferente de votos [3].

Para poder empezar con el análisis de los índices de poder, primero es necesario definir algunos conceptos como el poder de un participante de un SVP que intuitivamente puede entenderse como la capacidad de dicho participante de influir en una decisión [3]. Para medir esta capacidad surgen los índices de poder, que son teorías que bajo diferentes hipótesis de partida permiten estimar numéricamente el poder de cada participante. Entre todos los existentes, se ha decidido estudiar el índice de Banzhaf (IB), de Shapley-Shubik (ISS) y de Borda (IBR), pues son los más utilizados y con mayor teoría desarrollada a su alrededor. En particular, los dos primeros son los que mejor reflejan el concepto de poder, también dan una valoración más compleja que los votos obtenidos, mientras que el tercero simplemente se basa en diferentes vueltas, o rondas de votaciones, ponderando los votos obtenidos en cada vuelta con un peso en particular.

En los dos índices principales, IB y ISS, se tienen en cuenta las coaliciones que puedan formar los participantes pero, sin embargo, asumen que todas las coaliciones son equiprobables. Obviamente, esta asunción dentro del panorama político español es bastante fuerte y por tanto, durante el desarrollo de este TFG se decidió explorar estrategias que permitieran realizar una estimación de las probabilidades de las distintas coaliciones.

2.1. Análisis de redes sociales

Las redes sociales son una fuente de información enorme, pero hay que tener cuidado con las conclusiones que se obtienen de los datos recopilados de ellas, pues en función al contexto se sabe que los datos están mas sesgados, añadido a la topología de la red, pues es necesario conocerla para poder interpretar mejor los resultados [4].

2.1.1. Twitter

Actualmente, Twitter es una fuente de información muy importante. Muchas empresas sacan conclusiones y resultados basados en él. Por ejemplo, Dattio [5], o DragoSolutions [6], son empresas que obtienen información de Twitter con diferentes fines. En particular, los partidos políticos lo utilizan para hacer campaña electoral gratuita, así como informar a sus seguidores de mítines, propuestas, etc.

Para obtener los datos de Twitter se ha usado la API de Twitter que obtiene los seguidores de un usuario [7]. Con estos datos se ha intentado calcular las probabilidades de coalición en función a los seguidores de los partidos políticos y sus representantes.

Pero toda esta teoría tiene un inconveniente, el desconocimiento de la topología de la red. También hay que sumarle el sesgo de los usuarios de Twitter que muestran alguna preferencia política frente a la realidad [8]. Por tanto, las conclusiones obtenidas estarán influenciadas por factores externos, no sirve analizar simplemente datos de Twitter, sino que hay que interpretarlos. Con respecto al conocimiento de la topología se sabe que la gente que está relacionada tiende a tener ideas similares, y unos influencian a otros [4], por tanto, si un partido político tiene muchos seguidores habría que mirar la similitud y relación entre esos seguidores, porque quizá un fiel seguidor del partido X influencia a sus seguidores, inculcando sus ideas a estos, y finalmente consiguiendo que sus amigos sigan al partido X, cuando en realidad son seguidores del partido Y. Por tanto, las conclusiones no serían del todo correctas. Pues a pesar de que un usuario siga a X e Y, no quiere decir que la probabilidad de pacto entre X e Y sea alta.

2.1.2. Probabilidades de coalición

Para empezar con el estudio hay que definir cuales serán los usuarios relevantes:

- Mariano Rajoy (MR)
- Pedro Sánchez (PS)
- Pablo Iglesias (PI)
- Albert Rivera (AR)
- PP
- PSOE
- PODEMOS (P)
- CIUDADANOS (CS)

Para estimar la probabilidad de coalición entre partidos se calcularán una serie de probabilidades condicionadas que pretenden determinar, por ejemplo cuántos usuarios que siguen al PP siguen al PSOE, o cuántos usuarios que siguen a Pablo Iglesias siguen también a Ciudadanos, pues en función a estos valores se podría determinar la similitud entre partidos políticos. Para realizar este análisis se ha llevado a cabo un estudio con probabilidades condicionadas para lo que se han generado tres matrices, A, B y C, definidas a continuación.

La matriz A representa la probabilidad de seguir a un partido político dado que se sigue a determinado partido político.

$$A = \begin{pmatrix} P(PP|PP) & P(PSOE|PP) & P(P|PP) & P(CS|PP) \\ P(PP|PSOE) & P(PSOE|PSOE) & P(P|PSOE) & P(CS|PSOE) \\ P(PP|P) & P(PSOE|P) & P(P|P) & P(CS|P) \\ P(PP|CS) & P(PSOE|CS) & P(P|CS) & P(CS|CS) \end{pmatrix},$$

donde $P(X|Y)$ significa la probabilidad de que una cuenta de Twitter siga al partido X sabiendo que sigue al partido Y. Se puede apreciar que es una matriz que presenta el valor uno en la diagonal principal, pues $P(X|X) = 1$. Observar que la matriz no es simétrica puesto que, claramente, $P(X|Y) \neq P(Y|X)$.

La segunda matriz B representa la probabilidad de seguir a un líder político dado que se sigue a determinado partido político:

$$B = \begin{pmatrix} P(MR|PP) & P(PS|PP) & P(PI|PP) & P(AR|PP) \\ P(MR|PSOE) & P(PS|PSOE) & P(PI|PSOE) & P(AR|PSOE) \\ P(MR|P) & P(PS|P) & P(PI|P) & P(AR|P) \\ P(MR|CS) & P(PS|CS) & P(PI|CS) & P(AR|CS) \end{pmatrix}.$$

Por último, la tercera, y última matriz C, representa la probabilidad de seguir a un partido dado que se sigue a determinado líder político:

$$C = \begin{pmatrix} P(PP|MR) & P(PSOE|MR) & P(P|MR) & P(CS|MR) \\ P(PP|PS) & P(PSOE|PS) & P(P|PS) & P(CS|PS) \\ P(PP|PI) & P(PSOE|PI) & P(P|PI) & P(CS|PI) \\ P(PP|AR) & P(PSOE|AR) & P(P|AR) & P(CS|AR) \end{pmatrix}.$$

2.2. Índices de poder

En esta sección se presentarán los conceptos básicos en torno a sistemas de votación ponderados así como los fundamentos de los tres índices de poder considerados en este trabajo: Borda, Banzhaf y Shapley-Shubik.

2.2.1. Sistemas de votación ponderados

Las distintas teorías estudiadas en este TFG para el análisis del índice de poder de los partidos políticos se basan en el concepto de sistema de votación ponderado. Un sistema de votación ponderado (SVP) es un procedimiento de toma de decisiones en el que los participantes tienen un número distinto de votos [3].

Para este trabajo sólo se han considerado SVP con mayoría simple o mayoría absoluta así como votos de tipo sí, no y abstención.

En un sistema de votación ponderado tenemos n participantes, los cuales tienen votos que tendrán un valor de sí, no o abstención. En el caso de la política estos votos se conocen como escaños, que son calculados en función a los votos obtenidos en las elecciones. También podría ser un juego online, con grupos donde cada miembro del grupo tiene un determinado número de puntos a su favor. La clave es superar la cuota, q, que puede estar definida de dos formas: mayoría simple o mayoría absoluta. Es común suponer que todos los miembros de un partido político votarán lo mismo ante una decisión y, por tanto, los escaños obtenidos son equivalentes a los votos que tiene un partido ante una decisión.

Se definen formalmente los conceptos de mayoría simple y absoluta.

Definición **2.1.** *Sea S un SVP configurado con mayoría simple. Sea A el número de votos a favor de una medida y sea B el números de votos en contra.*

Entonces, si $A > B$ se considera que esa votación ha superado la cuota y, por tanto, la medida es aprobada.

En estos sistemas los votos pueden ser del tipo sí/no/abstención.

Observacion 1. *El número de votos de tipo abstención no se tienen en cuenta, ya que sólo se necesita tener mayor cantidad de votos a favor que en contra de la medida.*

Definición **2.2.** *Sea S un SVP configurado con mayoría absoluta. Sea m el número de votos en el sistema.*

Se define la cuota q como:

$$q = \left\lfloor \frac{m}{2} \right\rfloor + 1$$

Si el número de sí supera o iguala q, la medida es aprobada.

En estos sistemas los votos pueden ser del tipo sí/no.

Se muestra un ejemplo de cálculo de cuota en el contexto de la investidura del Presidente de Gobierno, suponiendo que los participantes fueran los cuatro partidos políticos con mayor número de escaños, es decir, Partido Popular (PP), Partido Socialista Obrero Español (PSOE), Podemos y Ciudadanos:

Ejemplo **2.1.** *El número de diputados obtenidos por el PP, PSOE, Podemos y Ciudadanos son: 137, 85, 71, 32, respectivamente.*

En el caso de este ejemplo primero se vota con mayoría absoluta. La cuota con mayoría absoluta es:

$$\left\lfloor \frac{137 + 85 + 71 + 32}{2} \right\rfloor + 1 = 163 \tag{2.1}$$

Si los diputados que votan sí superan, o igualan, 163 en número de votos, el Presidente es elegido, en caso contrario, se realiza una nueva votación con mayoría simple.

Supongamos que no se ha superado dicha cuota. Vamos a crear un escenario ficticio y a analizarlo con mayoría simple.

Sea un escenario de votación en el que PP vota sí, Ciudadanos se abstiene y Podemos vota no. En este caso, si PSOE votase sí, tendríamos que la cuota en este contexto sería de 71*, pues tenemos que superar el número de diputados que han votado no. Por tanto, el Presidente sería elegido, pues contaría con* $137 + 85 = 222$ *votos a favor, superando los* 71 *votos en contra de Podemos. Mientras que si PSOE votase que no, la cuota sería* 156*, pues tendríamos* 85 *(votos de tipo no del PSOE)* $+71$ *(votos de tipo no de Podemos)* $= 156$ *votos en contra y, por tanto, como votos a favor sólo hay* 137 *de los diputados del PP, el Presidente no saldría investido.*

También es necesario definir coalición.

Definición **2.3.** *Se denomina coalición a un subconjunto de los n participantes en un SVP que puede hacer que una medida sea aprobada o rechazada.*

Observacion 2. *En las agrupaciones o coaliciones se tienen en cuenta el poder de las mismas tanto para aprobar como para bloquear una medida, por ello, se suele diferenciar entre coalición ganadora que es aquella que permite a la ley adquirir el estado que la coalición ha votado, y coalición de bloqueo que es aquella que cuenta con la cantidad suficiente de votos para bloquear la ley e impedir que sea aprobada.*

Esta teoría está analizada desde el punto de vista de partidos electorales, pero es extrapolable a cualquier sistema de votaciones ponderado.

2.2.2. Índice de poder de Borda

Este es el índice propuesto por el matemático Jean-Charles de Borda en 1770. En realidad no es considerado un índice de poder sino un sistema de recuento de votos [9], pero en este trabajo se ha tratado como índice de poder ya que este recuento genera un ranking y por tanto, se puede considerar que el valor obtenido es el poder de cada participante. Es el índice de poder más simple ya que su teoría está basada en un sistema de voto ponderado preferencial, obteniendo dichas preferencias en diferentes vueltas de una votación. También utiliza pesos para cada una de esas vueltas que reflejan la importancia de cada vuelta, siendo la primera la más importante.

Como ejemplo, en [10] se propone utilizar el índice de Borda para la postulación de magistrados al Tribunal Supremo Electoral (TSE). Esto es propuesto por la Asociación de Investigación y Estudios Sociales (ASIES) [11]. Por otra parte, el sistema de votación de Francia utiliza un método de votación para elegir el Gobierno similar al propuesto. Se realizan varias rondas de votaciones, y en cada ronda se van eliminando los participantes que quedan en la cola [12].

Dado que en España no se sigue este método, se han empleado datos del Centro de Investigaciones Sociológicas (CIS) [13]. Estos muestran información estadística de cuáles son la primera y segunda opción para gobernar y qué partidos no deberían gobernar bajo ningún concepto según el criterio de los votantes.

Para este trabajo se ha decidido dar un peso de 10 a los votos obtenidos como primera opción, siendo este el doble del peso de los votos obtenidos como segunda opción, y 5 en segunda. Se ha decidido así para que los votos obtenidos como primera opción aporten el doble de poder que los votos obtenidos en segunda. Por último, -1 en el caso de obtener votos con la opción de no gobernar bajo ningún concepto, ya que más que dar poder, debería quitarlo.

2.2.3. Índice de poder de Banzhaf

Este índice es el más importante de este trabajo ya que es el que más teoría desarrollada a su alrededor tiene y porque permite hacer un análisis más completo de las votaciones con mayoría simple y absoluta. También a razón de que es el más citado en los fallos jurídicos, quizás debido a que Banzhaf llevo a juicio varios casos y sigue presentando expedientes a los Tribunales que evalúan SVP [3].

El índice de poder de Banzhaf fue introducido por el profesor de derecho de la Universidad de George Washington, John F. Banzhaf II en 1965 [14]. A pesar de esto, el verdadero creador de dicho índice fue Lionel Sharples Penrose en 1946. Por ello, a veces también se le llama índice Penrose–Banzhaf [15].

Siguiendo la notación en [3], la configuración de un SVP con cuota q puede escribirse como: $[q; x_1, ..., x_n]$ donde los x_i son los votos de cada participante en el sistema de votación ponderado.

Se procede a una explicación detallada y a la introducción de definiciones necesarias para la comprensión del índice. Lo primero es definir **votante basculante** y **hombre de paja**

Definición **2.4.** *Se denomina votante basculante al miembro de una coalición ganadora cuyo voto es esencial para que gane dicha coalición, o miembro de una coalición de bloqueo cuyo voto es esencial para que bloquee la misma.*

Observacion 3. *Se puede restringir el estudio a las coaliciones ganadoras pues las coaliciones de bloqueo serían simétricas cambiando el papel que juegan los votos sí por no.*

A continuación se define **hombre de paja**.

Definición **2.5.** *Se denomina hombre de paja al participante que no tiene ningún poder en un SVP. Es decir, no es basculante en ninguna coalición.*

Es interesante descubrir en una votación de un SVP un hombre de paja, pues su inclusión en alguna coalición no supone ningún cambio en términos de que la medida sea aprobada o bloqueada.

Se muestra un ejemplo con un hombre de paja para la mejor comprensión del concepto.

Ejemplo **2.2.** *Supongamos una situación de 3 participantes, A, B y C, con configuración de mayoría absoluta. A cuenta con 5 votos, B con 3 y C con 1.*

La cuota sería: 5

- *Si A vota sí la medida será aprobada, independientemente de los votos de B y C.*
- *Si A vota no la medida será bloqueada, independientemente de los votos de B y C.*

Por tanto, en este ejemplo B y C son hombres de paja ya que sus votos no son necesarios para aprobar/bloquear una medida.

Con todas las definiciones y observaciones anteriores, ya se puede explicar el índice de Banzhaf: métrica que contabiliza en cuántas votaciones de un SVP un participante es basculante. Se muestra con un ejemplo.

Ejemplo **2.3.** *Supongamos que tenemos un SVP con mayoría absoluta, con la siguiente configuración: [3; 2,1,1]*

Lo primero es comprobar que el valor de la cuota es correcto: $q = \left\lfloor \frac{2+1+1}{2} \right\rfloor + 1 = 3$

Veamos las $8 = 2^3$ *configuraciones posibles de este SVP:*

Combinaciones			**Votos**	**Estado de la ley**
SÍ	SÍ	SÍ	4	Aprobada
SÍ	SÍ	NO	3	Aprobada
SÍ	NO	SÍ	3	Aprobada
SÍ	NO	NO	2	Denegada
NO	SÍ	SÍ	2	Denegada
NO	SÍ	NO	1	Denegada
NO	NO	SÍ	1	Denegada
NO	NO	NO	0	Denegada

Tabla 2.1: Combinaciones de votos del ejemplo 2.2

Veamos en qué casos A es basculante, con el resto de participantes se haría de forma análoga. En la primera combinación (SÍ SÍ SÍ) se tienen 4 votos a favor pero si A se pasa al bando opositor, tendríamos NO SÍ SÍ, es decir, sólo se tendrían 2 votos a favor de sacar adelante la medida, lo que es menor que el valor de la cuota. Por tanto, la medida no saldría adelante, luego A es basculante en esta primera configuración, pues su voto es decisivo. En las dos siguientes combinaciones la ley también es aprobada y si A cambia de opinión la ley es denegada, pues en el segundo caso, el cambio de opinión de A conllevaría a que la propuesta tuviera sólo 1 voto a favor y pasara a ser bloqueada. Lo mismo sucede con la tercera de las configuraciones. Por tanto, A es basculante en tres de las configuraciones en las que su cambio de opinión en *coaliciones inicialmente ganadoras hace que las medidas pasen a ser denegadas. Luego, A tiene un poder por ahora de 3, pues en 3 combinaciones es basculante.*

Como ya hemos dicho previamente, no es necesario calcular las combinaciones donde A vota NO, pues es el caso simétrico sustituyendo SÍ por NO. Luego, obtendremos un poder de 3 también.

En resumen, el índice de poder de Banzhaf para A en este SVP sería de 6. Realizando los mismos cálculos para B y C se obtendría un índice de poder de 2 en ambos casos.

Solo quedaría normalizar, que por cuestiones que se abordaran en 2.3.2, entre el número de casos posibles, es decir, $2^3 = 8$.

Quedando finalmente: (75 %, 25 %, 25 %), que suma más del 100 % porque se da el caso que en varias votaciones más de un participante es basculante, por ejemplo, en la combinación dos tanto A como B son basculantes.

Como podemos ver en el ejemplo 2.3 este cálculo crece exponencialmente, pues hay que hacer 2^n combinaciones y después evaluar la mitad de ellas, comprobando en cuáles es basculante cada uno de los participantes.

El ejemplo 2.3 puede incitar a pensar que tener mismo poder y mismo número de votos es equivalente, pero no es cierto. En el ejemplo 2.4 se muestra como dos votantes con diferente número de votos tienen el mismo poder.

Ejemplo **2.4.** *Sea la siguiente configuración de un SVP: [51; 40, 30, 20, 10] con mayoría absoluta. Realizando el mismo proceso que en el ejemplo anterior, se obtienen los siguientes valores de IB: (5, 3, 3, 1). Normalizados quedan (31.25 %, 18.75 %, 18.75 %, 6.25 %).*

En este segundo ejemplo se aprecia lo contrario. En primer lugar que la suma es inferior al 100 % debido a que en muchas votaciones ninguno es basculante. En segundo lugar que B y C tienen el mismo poder a pesar de contar con distinto número de votos (30 y 20 respectivamente). Por tanto, queda claro que no es directamente proporcional el número de votos con el poder que tiene cada participante.

2.2.4. Índice de poder de Shapley-Shubik

Este es el índice propuesto por el matemático y economista Lloyd Shapley y el economista Martin Shubik en 1954 [3]. Fue el primer índice numérico ampliamente aceptado para evaluar el poder en los sistemas de votación. Su índice fue derivado del valor Shapley [16]. Este valor es un concepto fundamental en la teoría de Juegos [17], que como se mostrará en el siguiente apartado, un SVP es un Juego (3,2).

Para este índice es necesario manejar permutaciones, pues es la clave principal para esta métrica. Una permutación es una variación del orden de un conjunto sin repeticiones, y cada variación debe ser considerada para determinar el índice de Shapley-Shubik. Debido a esto, el coste computacional será bastante alto pues en un SVP con n participantes el número de permutaciones será de $n!$ ya que para hacer las permutaciones es necesario hacer lo siguiente:

- Para el primer elemento, n votantes posibles
- Para el segundo, n-1
- ...
- Para el último, un único votante

Luego se tiene $n \times (n-1) \times ... \times 1 = n!$. Claramente cuando n es pequeño el coste es asumible, pero a medida que n crece, el coste computacional se hace demasiado grande.

En el panorama actual de la política española se cuenta con nueve partidos políticos, unificando los distintos grupos de Podemos según indica [18], por lo que el número de permutaciones a analizar sería de 362880, que es un número elevado de casos a tener en cuenta, pero el índice de poder es calculable. Sin embargo, en años pasados, por ejemplo en 1979, el número de partidos era de 14, por lo que la cuantía de permutaciones a generar era de 87178291200, por lo que resulta imposible calcular el índice de poder, pues ni siquiera es posible generar todas las permutaciones.

Se puede considerar que cada una de las permutaciones representa un espectro de opinión sobre un asunto, en el que el primer votante de la permutación es el más convencido, el segundo un poco menos y así sucesivamente, hasta llegar al último, que considera odiosa dicha medida.

Una vez está definido el objeto que se utilizará, las permutaciones, se procede a definir los conceptos necesarios.

Definición **2.6.** *Se denomina* ***votante pivotante*** *de la permutación a aquel votante que se encuentra en la posición que convierte a la coalición en ganadora.*

Observacion 4. *En todas las permutaciones hay siempre un votante pivotante.*

Para entenderlo mejor se muestra un ejemplo.

Ejemplo **2.5.** *Sea un SVP [3;2,1,1] con mayoría absoluta y cuota $q = 3$.*

Se tienen tres participantes A, B y C con votos (2,1,1).

Se muestran las posibles permutaciones:

1. *(A B C), sus votos serían (2 1 1). A no sería pivotante, pues sólo con sus votos la medida no sería aprobada. En cambio, si se añade B a la coalición los votos a favor serían 3, que igualan la cuota, por lo que la medida sería aprobada. Por tanto, B sería el votante pivotante de esta permutación.*

2. *(A C B), sus votos serían (2 1 1). A no sería pivotante, pero del mismo modo que anteriormente, si C se incluye en la coalición la medida sería aprobada. Por lo que C es pivotante en esta permutación.*

3. *(B A C), sus votos serían (1 2 1). A sería el votante pivotante de esta permutación.*

4. *(B C A), sus votos serían (1 1 2). B no sería pivotante y si se integra C a la coalición la medida tampoco sería aprobada pues sumarían 2 votos a favor. Finalmente, cuando se une A a la coalición la medida es aprobada. Por tanto, A es el votante pivotante de esta permutación.*

5. *(C A B), sus votos serían (1 2 1). Este es el mismo caso que en 3 solo que C juega el papel de B. Luego A es el votante pivotante.*

6. *(C B A), sus votos serían (1 1 2). Este es el mismo caso que 4 jugando C el papel de B y viceversa. Luego A es el votante pivotante.*

Resumiendo, A es pivotante en 4 permutaciones, B en 1 y C en 1.

Hay que destacar que en el índice de poder de Shapley-Shubik tener mismo número de votos implica tener mismo poder, pues las permutaciones en las que los participantes x_i serán pivotantes son las mismas. Esto se debe a que las permutaciones cambiarán conceptualmente pues una será (A B ...) y la otra (A C ...), pero si el número de votos de B y C es el mismo la coalición AC y AB sumarán los mismos votos y si B era pivotante, C también lo será, pues si AB (suma de los votos de A más los votos de B) es mayor que q, entonces AC (suma de los votos de A más los votos de C) también será mayor que q.

Con todo esto, ya se puede introducir la definición de dicho índice.

Definición **2.7.** *Se define el índice de poder de Shapley-Shubik de un votante como el porcentaje de las permutaciones en las que dicho votante es pivotante.*

Volviendo al ejemplo 2.5 el ISS quedaría de la siguiente forma:

- A era pivotante en 4 permutaciones, por tanto su índice de Shapley-Shubik será $\frac{4}{3!} = \frac{2}{3} = 66,67\,\%$
- B era pivotante en 1 permutación, por tanto su índice de Shapley-Shubik será $\frac{1}{3!} = \frac{1}{6} = 16,67\,\%$
- C era pivotante en 1 permutación, por tanto su índice de Shapley-Shubik será $\frac{1}{3!} = \frac{1}{6} = 16,67\,\%$

Claramente, como se ha dicho en la observación 4, la suma de todos los índices es el $100\,\%$ ($66,67\,\% + 16,67\,\% + 16,67\,\% = 100\,\%$), pues en todas las permutaciones hay un único votante pivotante (existencia y unicidad).

Un caso particular de un SVP en el que el cálculo del ISS es fácil, es aquel en el que todos los participantes cuentan con el mismo número de votos, dado que todos serán pivotantes en el mismo número de permutaciones. Si se tienen n participantes, cada uno con $i\,\%$ poder, pues todos tienen el mismo poder, sabiendo que $i\,\% + i\,\% + ... + i\,\% = n \times i\,\% = 100\,\%$, entonces $i = \frac{100}{n \times 100} = \frac{1}{n}$.

Otro caso particular un poco más complejo, pero fácil de calcular, es un SVP en el que todos los participantes tienen el mismo número de votos salvo uno. En este caso se puede evitar calcular muchas permutaciones, ya que los $n-1$ participantes con el mismo número de votos son equivalentes y sus permutaciones serán iguales, sólo cambiará el nombre del participante, pero el poder que refleja es el mismo.

Al intentar aplicar este índice a la vida real surgen casos extremos. Como por ejemplo, el Colegio Electoral de Estados Unidos, que cuenta con 50 estados más el Distrito de Columbia, formando un total de 51! permutaciones [3].

El primer problema que plantea este caso es calcular 51!, para lo que se propone aproximar los factoriales usando la fórmula de Stirling [19]:

$$n! \approx \sqrt{2\pi n} n^n e^{-n} \tag{2.2}$$

Existen algoritmos numéricos que aproximan el valor de factoriales de números grandes, como por ejemplo PrimeSwing [20]. Para comprobar una aproximación de 51! se ha recurrido a una aplicación web [21]. El valor obtenido tiene 68 dígitos, por lo que es inviable enumerar tal número de permutaciones. Para dar solución a casos similares, en el artículo [22] se muestran algoritmos para el cálculo del índice de poder de Shapley-Shubik cuando el número de participantes es relativamente grande. John P. Lambert también propuso soluciones a este problema en [23]. Sin embargo, en este trabajo no se van a abordar dichos algoritmos pues se extralimitan de los objetivos del mismo.

2.2.5. Comparación entre IB y ISS

Decidir qué índice describe mejor la distribución de poder en un SVP es subjetivo. La clave es distinguir permutaciones y combinaciones. Una permutación representa la variedad de opinión sobre un asunto, por lo que ISS es más acertado si hay un amplio conjunto de opiniones sobre la mayoría de cuestiones sobre las que decidir, mientras que IB es preferible si la pregunta no admite una amplia gama de opiniones [3].

2.3. Teoría de Juegos (3,2)

Este trabajo se ha centrado en estudiar las distintas votaciones al Congreso de los Diputados desde el año 1977. Cada una de estas votaciones puede ser interpretada como un juego en el que cada participante tiene tres opciones: votar sí, votar no o abstenerse; y la medida puede tener dos estados: aprobada o rechazada. Por ello, se denominarán juegos (3,2) de ahora en adelante a los SVP. La gran aportación de esta teoría de juegos es el cambio a mayoría simple. En los SVP anteriores sólo se tenían en cuenta configuraciones de mayoría absoluta sin abstenciones. Gracias a esta teoría se permite pasar a un SVP con abstenciones y configuraciones con mayoría simple. Toda la teoría que se va a describir a continuación se basa en [24]. A pesar de poder aplicar esta teoría al resto de índices, se ha decidido aplicarla sólo al índice de poder de Banzhaf, ya que este índice de poder es el principal de este trabajo.

En estos juegos existen 3 grupos: los miembros que votan sí (A_1), los participantes que votan abstención (A_2) y los que votan no (A_3). Los votantes pueden cambiar de grupo, es decir, pueden cambiar su voto. A estos procesos se les denominará cambios, que pueden ser cambios a corto o cambios a largo. A continuación se define cada uno de ellos.

Definición **2.8.** *Se llama cambio a corto cuando un votante que se encuentra en el grupo A_i pasa al grupo A_{i+1} o al grupo A_{i-1}. Y se representa de la siguiente forma: $A_{\downarrow i}$ o $A_{\uparrow i}$, respectivamente.*

Observacion 5. *Si el participante $x \in A_1$, entonces $A_{\uparrow 1} = A_0$ no tiene sentido. Este participante sólo podrá realizar un cambio a corto de la forma $A_{\downarrow 1} = A_2$. De forma análoga pasa si $x \in A_3$, que sólo podrá realizar $A_{\uparrow 3} = A_2$.*

Definición **2.9.** *Se denomina cambio a largo cuando un votante que se encuentra en el grupo A_i pasa al grupo A_{i+2} o al grupo A_{i-2}. Se representa de la siguiente forma: $A_{\downarrow\downarrow i}$ o $A_{\uparrow\uparrow i}$, respectivamente.*

Observacion 6. *Sólo tiene sentido hablar de cambios a largo cuando se tiene $x \in A_1$ o $x \in A_3$, pues en el resto de casos no están definidos.*

Un juego (3,2) puede ser cualquier otro sistema en el que los participantes tengan 3 opciones a elegir en una decisión y que el resultado sean dos características de la misma. Por ejemplo, en un juego online cada participante puede elegir atacar, sanar o no hacer nada; podemos hacer una simple asociación: atacar equivaldría a sí, sanar a no y no hacer nada a abstención. El resultado, en lugar de ser aprobar o rechazar una medida, puede ser atacar o sanar también. Con esto se puede medir el poder de decisión que tiene cada jugador en esa partida.

Con esta interpretación se puede hacer un ranking de los usuarios más poderosos, o mejor dicho, con mayor poder de decisión, entendidos como aquellos que son basculantes en el mayor número de combinaciones.

A continuación se diferencian dos configuraciones distintas de los juegos: sin abstención, juegos (2,2); y con abstención, juegos(3,2).

2.3.1. Juegos sin abstención

Estos son los juegos que corresponden a SVP configurados con mayoría absoluta y en los que todos los participantes tienen que votar sí o no. Como ya es sabido, para aprobar una medida se tiene que obtener más votos a favor que en contra.

Se procede a analizar estos juegos para obtener una fórmula de la cuota q del juego (2,2).

- Se llamará s al número de votos a favor de una medida y será un número entero, dado que representa un número de votantes.
- Se llamará d al número de votos en contra de una medida y será un número entero, dado que representa un número de votantes.
- Se llamará n al número de participantes individuales en el juego y también será un número entero.

Pese a que n puede entenderse como el número de partidos políticos, en este estudio de la cuota no es así. Este número es el total de votantes del juego. Ahora bien, como estos votantes pertenecen a un partido político, los votos de los m participantes del juego (2, 2) pertenecientes a un mismo partido será del mismo tipo (sí/no), pero cada uno de esos m participantes están presentes de forma individual en el juego (n votos en total).

Primero se supondrá que n es par. En el caso peor, como sólo se tienen dos opciones en la votación y se necesitan al menos la mitad de votos totales, se obtiene esta primera ecuación:

$$s \geq \frac{n}{2}, \text{ si n es par} \tag{2.3}$$

Para obtener d, basta con calcular d como $d = n - s$, pues todos los participantes votan sí o no. Usando la ecuación 2.3, se tiene $d = n - \frac{n}{2} = \frac{n}{2}$, por lo que $d = s$.

Según está definido en los SVP, igualar en número los votos a favor y los votos en contra no aprueba la medida: el número de votos a favor debe ser estrictamente mayor al número de votos en contra. Por lo que se necesita que:

$$s \geq \frac{n}{2} + 1 \tag{2.4}$$

.

Pero si n fuese impar, la ecuación 2.3 no tendría sentido, pues s es un valor no entero, en contra de su naturaleza. Por lo que es necesario aplicar el suelo, $\lfloor\rfloor$ (quitar decimales), o el techo, $\lceil\rceil$ (aumentar en uno el valor entero del número racional y despreciar los decimales) sobre $\frac{n}{2}$ para trabajar con números enteros. Si se utilizase el techo, se tendría:

$$s \geq \left\lceil \frac{n}{2} \right\rceil \tag{2.5}$$

Quedando por tanto la fórmula 2.4 si n es par y 2.5 si n fuese impar, lo cual es poco práctico.

Si por el contrario, se escogiera la operación suelo, se seguiría que $s \geq \lfloor \frac{n}{2} \rfloor$, y puesto que todos los participantes votan sí o no, $d = n - s = n - \lfloor \frac{n}{2} \rfloor = \lceil \frac{n}{2} \rceil$. Por ello, d sería mayor que s y la medida no sería aprobada. Este caso es muy parecido al caso de n par; es equivalente a empatar en números de votos a favor y en contra, por tanto se puede plantear la misma solución que en 2.4:

$$s \geq \left\lfloor \frac{n}{2} \right\rfloor + 1, \text{ independientemente de n par o impar} \tag{2.6}$$

Quedando entonces la fórmula que se ha definido en secciones previas como la cuota en SVP con mayoría absoluta. Este análisis ha sido desarrollo propio de este trabajo para comprobar las fórmulas descritas sobre la cuota en secciones previas.

Trivialmente, en estos juegos $A_2 = \emptyset$ ya que no existe la opción de abstenerse. Obviamente, $juego(2, 2) \subset juego(3, 2)$.

Además, al tener $A_2 = \emptyset$, sólo tendrán sentido los cambios a largo, es decir, cambios de sí a no y no a sí ($A_{\downarrow\downarrow 1}$ o $A_{\uparrow\uparrow 3}$).

Como resumen, se tienen dos posibilidades: votar sí, por lo que el participante $p \in A_1$; o votar no, por lo que el participante $p \in A_3$. Se tendrá que analizar cada una de las combinaciones, realizando cambios a largo para cada participante p:

1. Si $p \in A_1$, entonces se realizará un cambio a largo $A_{\downarrow\downarrow 1}$ votando que no y se analizará si p es basculante.

2. Si $p \in A_3$, entonces se realizará un cambio a largo $A_{\uparrow\uparrow 3}$ votando que sí y se analizará si p es basculante.

Toda esta teoría está desarrollada para calcular IB, pues el ISS se centra en permutaciones que no tienen en cuenta cambios a corto y a largo. Estos juegos son casos en los que no se tiene una amplia gama de opiniones, tan solo 2. Por tanto, es preferible IB en SVP, tal y como se describió en el apartado 2.2.5.

2.3.2. Juegos con abstención

Estos son los juegos que corresponden a SVP configurados con mayoría simple y en el que todos los participantes pueden votar sí, no o abstención (3 opciones de voto). Igual que en el apartado 2.3.1, para aprobar una medida se tienen que obtener más votos a favor que en contra. A diferencia del caso anterior, los Juegos (3,2) sí que pueden tener $A_2 \neq \emptyset$, por tanto, tiene sentido hablar de cambios a corto.

Observacion 7. *Las votaciones en las que $A_2 = \emptyset$, son casos particulares de juegos sin abstención. Por tanto, los juegos con abstención incluyen a los juegos sin abstención, como ya se ha mencionado en 2.3.1.*

Ahora la cuota será dinámica. Se calculará en cada una de las votaciones en particular y dependerá de lo que vote cada uno de los participantes, ya que sólo es necesario superar en número los votos de tipo sí a los de tipo no. Por ejemplo, si en un SVP de m participantes y n votos totales, hay un voto a favor y los restantes, $n - 1$, son votos en abstención, la medida saldrá aprobada. Para analizar estos juegos se introduce la siguiente notación:

- Se llamará a al número de votos en abstención y será un número entero, dado que representa un número de votantes.

- Se utilizarán s, d y n según lo descrito en la sección 2.3.1.

Para aprobar una medida será necesario que $s \geq d + 1$. En realidad, los participantes que se abstienen no están siendo representados a la hora de determinar si una ley es aprobada o no.

Su aportación sólo está siendo reflejada en no pertenecer al grupo A_3 (participantes que votan no). Podría plantearse un nuevo juego (3,2), en el que $A_2 = \emptyset$ con $m = s + d = n - a$ participantes. Esto permite suprimir casos a analizar, pues mientras que en estos juegos las posibilidades de voto son 3 (sí/no/abstención), es decir, las combinaciones de votaciones son 3^n, en los juegos sin abstención las combinaciones se reducen a 2^n, pues sólo tienen dos opciones de voto (sí/no). Por tanto, podemos concluir que un juego (3,2) es una combinación de juegos (2,2). Se detallara la teoría desarrollada sobre la combinación de juegos (2,2) para obtener un juego (3,2) en la sección 3.3.1.

El problema de juegos (3,2) puede enfocarse de otra manera teniendo en cuenta los resultados obtenidos en el artículo [24]. En este artículo se demuestra que el índice de poder de Banzhaf de un votante en un juego (3,2) puede evaluarse omitiendo los cambios del votante a corto, es decir, sólo es necesario evaluar el cambio de un voto sí a no (y viceversa) y comprobar si es basculante.

Antes de presentar la proposición de [24] que permitirá reducir el coste computacional de los juegos (3,2), es necesario definir las variables Bz_i y Bz_i':

1. Se denotará Bz_i al índice de poder de Banzhaf del usuario $i - \acute{e}simo$ en un juego en el que el participante $i - \acute{e}simo$ puede cambiar su voto entre sí y no (sólo tiene cambios a largo).
2. Se denotará Bz_i' al índice de poder de Banzhaf del usuario $i - \acute{e}simo$ en un juego en el que el participante $i - \acute{e}simo$ puede cambiar su voto entre sí, no y abstención (tiene cambios a corto y largo).

Se recuerda que para calcular el IB de un votante es necesario comprobar en cuántas combinaciones el votante es basculante. Luego, es un índice que se calcula en cada una de las votaciones posibles para cada uno de los participantes. Cambiando sus opciones de voto entre las opciones posibles y evaluando el resultado según lo explicado en 2.2.3.

Proposición 1. *Sea J un juego(3,2). Sea $X = \{1, 2, ..., n\}$ el conjunto de votantes cada uno con x_i votos, entonces si $i \in X$ se verifica que $Bz_i = Bz_i'$*

En este trabajo se denominará X al conjunto formado por los partidos políticos y cada x_i a los votos del partido $i - \acute{e}simo$.

Demostración de la proposición 1. Sea $|A|$ el cardinal de un conjunto A de coaliciones, Λ el conjunto de coaliciones ganadoras, y sea $\eta_i = |\{A : A \in \Lambda, A_{\downarrow i} \notin \Lambda\}|$ el número de coaliciones ganadoras.

Lo primero que se observa en esta definición de η_i es que:

$$|\{A : i \in A_2, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}| = |\{A : i \in A_1, A_{\downarrow i} \in \Lambda, A_{\downarrow\downarrow i} \notin \Lambda\}| \tag{2.7}$$

En la parte izquierda de la ecuación 2.7 se tienen estas votaciones: (..., abstención, ...) y la medida es aprobada. Si se cambia el voto a no quedan: **(..., no, ...)** y la medida es rechazada.

Si para cada una de las votaciones anteriores se tuviera la votación: (..., sí, ...) dónde solo cambia abstención por sí, la medida también sería aprobada pues contaría incluso con mayor número de votos a favor que en la votación anterior donde se abstenía el votante $i - \acute{e}simo$. Si ahora se cambia su voto a no, queda: **(..., no, ...)**, por lo que la medida también es rechazada. Por tanto, el votante $i - \acute{e}simo$ es basculante en esta votación. Justamente, esto es la parte derecha de la ecuación 2.7.

Por ello, se puede concluir que para cada votación en la que i escogía abstención y era basculante se tiene una votación equivalente en la que i vota sí y es también basculante.

Por tanto, se puede formular η_i teniendo en cuenta sólo los cambios a largo:

$$\eta_i = |\{A : A \in \Lambda, A_{\downarrow\downarrow i} \notin \Lambda\}| \tag{2.8}$$

.

Con esta nomenclatura se tiene la siguiente ecuación para el IB en un juego (3,2)

$$Bz_i = \frac{\eta_i}{3^{n-1}} \tag{2.9}$$

El denominador es $n-1$ porque el participante $i-ésimo$ tiene su voto fijo cuando se calcula el índice de poder sobre él, y cuando se analiza una votación particular es cuando se cambia el voto del participante de sí a no (o de no a sí) y evaluamos el estado resultante de la medida.

Como Bz_i' calcula el poder del participante $i-ésimo$ haciendo cambios a cortos y largos sin dejar al participante fijo se tienen $3 \times 3^{n-1}$ posibilidades (sí -> abstención, abstención -> no, sí -> no, o sus simétricos). Por tanto, $Bz_i' = \frac{\eta_i'}{3^n}$.

Se puede escribir η_i' como la suma de η_i^* y η_i^{**}, siendo estas:

$\eta_i^* = |\{A : i \in A_1, A \in \Lambda, A_{\downarrow\downarrow i} \notin \Lambda\}| + |\{A : i \in A_2, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}|$, vota sí y es basculante cambiando a abstención y a no.

$\eta_i^{**} = |\{A : i \in A_3, A \in \Lambda, A_{\uparrow\uparrow i} \notin \Lambda\}| + |\{A : i \in A_2, A \in \Lambda, A_{\uparrow i} \notin \Lambda\}|$, vota no y es basculante cambiando a abstención y a sí.

Por tanto, si se quiere ver $Bz_i = Bz_i'$, basta ver que $\eta_i' = 3\eta_i$.

Se reescriben η_i^* y η_i^{**} como sigue:

$\eta_i^* = \eta_i + |\{A : i \in A_2, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}|$, $\eta_i^{**} = \eta_i + |\{A : i \in A_2, A \in \Lambda, A_{\uparrow i} \notin \Lambda\}|$.

Con esta nueva notación se tiene:

$\eta_i' = 2\eta_i + |\{A : i \in A_2, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}| + |\{A : i \in A_2, A \in \Lambda, A_{\uparrow i} \notin \Lambda\}|$

Por lo que queda probar que: $\eta_i = |\{A : i \in A_2, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}| + |\{A : i \in A_2, A \in \Lambda, A_{\uparrow i} \notin \Lambda\}|$ para tener la igualdad buscada.

Pero justamente, estas expresiones de la derecha coinciden con:

$|\{A : i \in A_1, A_{\downarrow i} \in \Lambda, A_{\downarrow\downarrow i} \notin \Lambda\}| + |\{A : i \in A_1, A_{\downarrow i} \notin \Lambda\}| = |\{A : i \in A_1, A \in \Lambda, A_{\downarrow i} \notin \Lambda\}|$.

□

Se puede comprobar en la demostración que los valores del índice de poder de Banzhaf son normalizados entre el total de casos posibles (3^n) pues es una manera lógica de homogeneizar dichos índices de poder. Téngase en cuenta que con este tipo de normalización puede suceder que la suma de los índices de poder de los participantes sea superior a 1, esto es así porque podría darse el caso en el que en una misma votación dos participantes sean basculantes, como se muestra en el ejemplo . Del mismo modo, se pueden dar votaciones en las que ningún participante sea basculante, por lo que la suma de los índices de poder de los participantes sea inferior a 1, como se muestra en el ejemplo . Esta proposición permite pasar de 3^n votaciones distintas a 2^n, pues se reducen las opciones de voto de sí/no/abstención a sí/no. Esto, cuando n es grande, disminuye el tiempo computacional.

3

Diseño y desarrollo

3.1. Herramientas y tecnologías

En este capítulo se comentarán las herramientas desarrolladas, el conjunto de datos que han sido utilizados así como los algoritmos desarrollados para el cálculo de los índices de poder.

3.1.1. Herramienta en Shiny. Herramienta en Python

En este Trabajo de Fin de Grado se han implementado dos herramientas. La principal calcula los tres índices de poder: índice de poder Borda, índice de poder de Banzhaf e índice de poder de Shapley-Shubik; y la secundaria se utiliza para obtener probabilidades de coalición basadas en los datos de Twitter.

La aplicación principal encargada de calcular los índices de poder está implementada en R [25], pues es un lenguaje pensado para realizar análisis estadísticos y justamente ese es el objetivo de este trabajo. El entorno de desarrollo fue RStudio [26], y el framework Shiny [27], ya que ambos facilitan el desarrollo de aplicaciones en R. Este es un lenguaje nuevo que no se había enseñado durante los estudios del Doble Grado, por lo que ha sido necesaria mucha documentación para aprender a manejarlo como el cursillo detallado en [28] y [29].

La otra gran ventaja del lenguaje R es la facilidad con la que permite trabajar con combinaciones y permutaciones, elementos clave de los índices de poder de Banzhaf y Shapley-Shubik, respectivamente. Por tanto, fue una elección muy acertada.

Añadido a esto, se usó una librería especial para tener una interfaz simple e intuitiva, esta se denomina *shinydashboard* [27]. También se buscó una librería para trabajar con ficheros Excel [30], pero finalmente no fue necesaria.

La segunda aplicación para obtener los datos de Twitter está desarrollada en Python porque es más estable y consume menos recursos. Además, existe el paquete *Tweepy* [31], que encapsula todas las llamadas a la API REST de Twitter [7], siendo transparente para el usuario el tema de *timeout*, limitaciones de peticiones, etc.

Se usan conjuntos y Python aporta muchas funcionalidades para este tipo de datos, ya que tiene implementadas funciones que realizan la intersección, unión y calculan el cardinal de un conjunto, por lo que se decidió utilizar dicho lenguaje. Con la aplicación se generan los conjuntos de *followers* de cada partido político: PP, PSOE, Podemos y Ciudadanos; y de cada político: Mariano Rajoy, Pedro Sánchez, Pablo Iglesias y Albert Rivera. Posteriormente se calcula el número de *followers* que tienen cada uno, intersección entre ellos, y con todos estos datos se completan las matrices definidas en el apartado 2.1.2.

Se muestra en la Figura 3.1 la interfaz gráfica de la herramienta desarrollada en R Shiny.

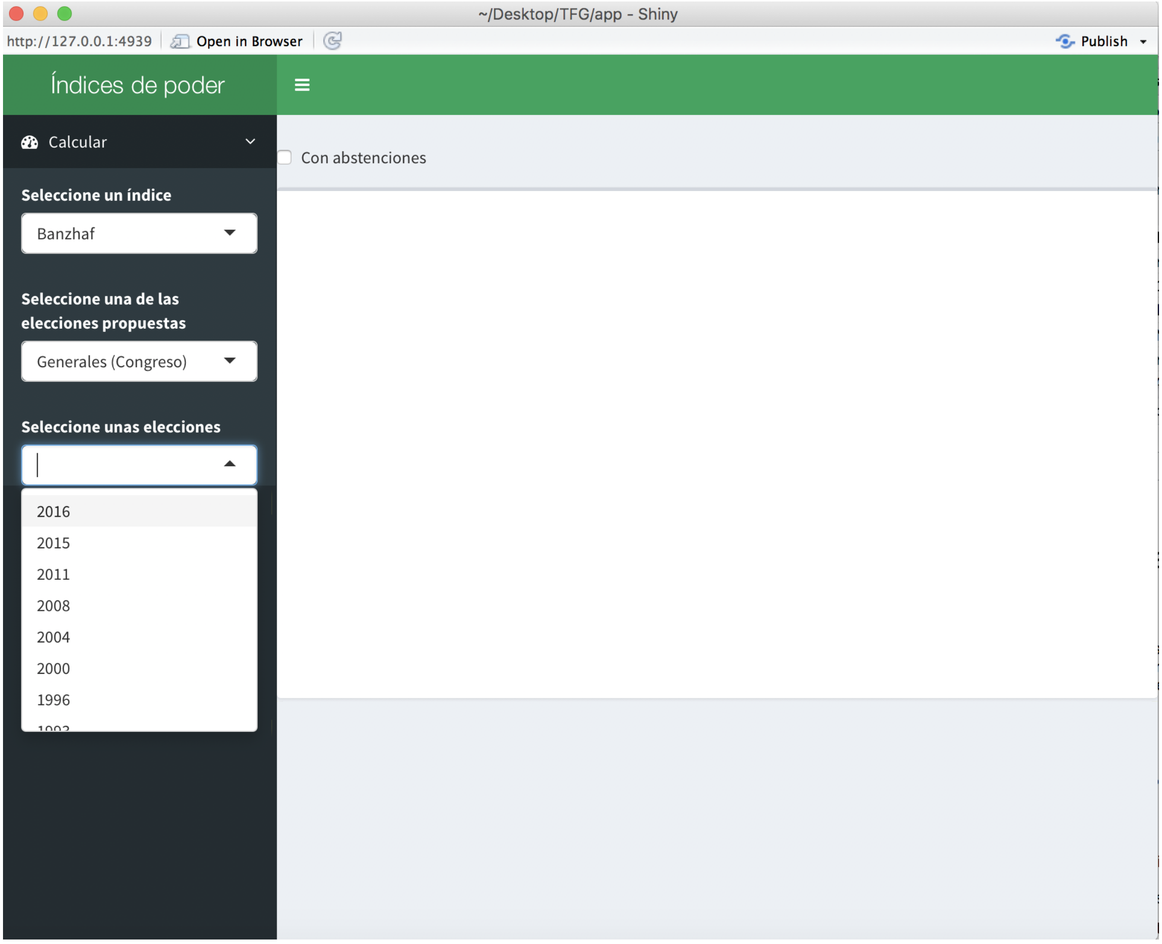

Figura 3.1: Interfaz gráfica de la aplicación de cálculo de los índices de poder.

En la parte izquierda presenta un panel de selección de índice de poder. Una vez se ha seleccionado un índice de poder, el índice de poder de Banzhaf en la Figura 3.1, permite elegir entre las elecciones del Congreso de los Diputados (opción seleccionada) o configuración libre. Por último, tiene una barra desplegable con cada una de las elecciones al Congreso de los Diputados. Si se pincha sobre una elección particular, la aplicación calcula el índice de poder y muestra histogramas como los que se detallan en la sección 4.2.

Se aprecia que es una herramienta muy simple, pues todo el trabajo desarrollado está en los algoritmos internos de la aplicación así como en la teoría estudiada para la comprensión de los índices de poder y teorías sobre los sistemas de votación que permiten la abstención.

3.2. El conjunto de datos

Los datos que se han dispuesto para la realización de este trabajo han sido obtenidos de tres fuentes: Twitter, Ministerio del Interior [2] y Centro de Investigaciones Sociológicas (CIS) [13].

Los datos de Twitter son ficheros de datos que contienen en cada línea un identificador que representa el ID de los *followers* de cada uno de los partidos y de los políticos. En total hay ocho ficheros por día: 4 de partidos (PP, PSOE, Podemos y Ciudadanos) y 4 de políticos (Mariano Rajoy, Pedro Sánchez, Pablo Iglesias y Albert Rivera).

Estos ficheros son tratados mediante el programa en Python, donde se generarán variables de tipo conjunto (en el sentido matemático) que contendrán los identificadores para cada uno de los partidos políticos: PP, PSOE, Podemos y Ciudadanos; y de cada político: Mariano Rajoy, Pedro Sánchez, Pablo Iglesias y Albert Rivera que facilitarán el cálculo de las probabilidades condicionadas gracias al tipo de variable *set* de Python, que permitirá obtener las intersecciones de dichos conjuntos cuyos elementos serán los identificadores de usuario. Con estos datos se completan las matrices definidas en el apartado 2.1.2.

Los datos obtenidos del CIS han sido recopilados de una encuesta preelectoral a las elecciones generales de 2016 [1]. De todas las cuestiones que incluye esta encuesta, se ha determinado que la relevante para el estudio del índice de Borda es la decimotercera pregunta, que se encuentra ubicada en la página 16 de dicha encuesta. En la Figura 3.2 se muestran los resultados obtenidos para esta pregunta.

	0 Con toda seguridad, no lo votaría nunca	1	2	3	4	5	6	7	8	9	10 Con toda seguridad, lo votaría siempre	N.S.	N.C.	**(N)**
PP	57,0	2,2	2,1	2,0	1,8	6,8	2,2	2,3	3,7	2,0	11,3	3,4	3,2	(17.488)
PSOE	43,4	3,1	4,6	4,7	4,4	11,7	3,8	3,5	4,5	2,0	7,1	4,0	3,3	(17.488)
Podemos	54,4	3,0	3,5	3,2	2,8	7,6	2,9	3,1	3,8	1,9	6,8	3,7	3,2	(17.488)
Ciudadanos	48,0	2,9	4,1	4,4	4,3	11,9	4,8	3,7	3,6	1,5	2,9	4,5	3,3	(17.488)

Figura 3.2: Tabla de la encuesta preelectoral de 2016 del CIS [1]

Esta pregunta refleja el porcentaje de:

1. Primera opción de los ciudadanos con respecto a cada partido. (Columna 10).
2. Segunda opción de los ciudadanos con respecto a cada partido. (Columna 5).
3. Última opción de los ciudadanos con respecto a cada partido. (Columna 0).

La elección de esta pregunta vino motivada porque es la que más se asemejaba a la idea que defiende el índice de Borda, pues sería lo equivalente a realizar tres vueltas. La décima columna representaría a quién se votaría en primera vuelta, como primera opción, y la quinta columna representaría a quién se votaría en segunda vuelta, como segunda opción. La última vuelta, más que una ronda de votación, es una ronda de eliminación en la que cada votante indicaría a qué partido no votaría bajo ningún concepto.

Por otro lado, los datos obtenidos del Ministerio del Interior son en formato Excel [2]. Lo primero destacable es que los archivos están dañados, lo que ha supuesto una carga adicional de trabajo, pues se ha tenido que cambiar la extensión de cada uno de los archivos, uno por cada una de las elecciones al Congreso de los Diputados. Además, la estructura de los mismos no era la más idónea para un tratamiento de datos rápido y sencillo. En la Figura 3.3 se muestra un ejemplo.

Consulta de Resultados Electorales

Congreso | Junio 2016

Total nacional

Población:	46.624.382	
Número de municipios:	8.125	
Número de mesas:	57.531	
Censo electoral sin CERA:	34.596.892	
Votantes a las 14:00 h :	12.755.388	36,87%
Votantes a las 18:00 h :	17.715.754	51,21%
Votantes a las 20:00 h :	24.157.982	69,83%
Censo CERA:	1.924.021	
Solicitudes voto CERA aceptadas:	169.658	
Votantes CERA:	121.277	6,30%
Total censo electoral:	36.520.913	
Total votantes:	24.279.259	66,48%
Abstención:	12.241.654	33,52%
Votos válidos:	24.053.755	99,07%
Votos nulos:	225.504	0,93%
Votos a candidaturas:	23.874.674	99,26%
Votos en blanco:	179.081	0,74%

Siglas	Candidatura	Votos	% válidos	% censo	% candidaturas	Diputados
PP	PARTIDO POPULAR	7.941.236	33,01%	21,74%	33,26%	137
PSOE	PARTIDO SOCIALISTA OBRERO ESPAÑOL	5.443.846	22,63%	14,91%	22,80%	85
PODEMOS-IU-EQUO	UNIDOS PODEMOS	3.227.123	13,42%	8,84%	13,52%	45
C's	CIUDADANOS-PARTIDO DE LA CIUDADANÍA	3.141.570	13,06%	8,60%	13,16%	32
ECP	EN COMÚ PODEM-GUANYEM EL CANVI	853.102	3,55%	2,34%	3,57%	12
PODEMOS-COM	COMPROMÍS-PODEMOS-EUPV: A LA VALENCIANA	659.771	2,74%	1,81%	2,76%	9
ERC-CATSÍ	ESQUERRA REPUBLICANA/CATALUNYA SÍ	632.234	2,63%	1,73%	2,65%	9
CDC	CONVERGÈNCIA DEMOCRÀTICA DE CATALUNYA	483.488	2,01%	1,32%	2,03%	8
PODEMOS-EN MAREA-ANOVA-EU	EN MAREA	347.542	1,44%	0,95%	1,46%	5
EAJ-PNV	EUZKO ALDERDI JELTZALEA-PARTIDO NACIONALISTA VASCO	287.014	1,19%	0,79%	1,20%	5

Figura 3.3: Ejemplo de fichero de datos de las elecciones al Congreso obtenido de [2]

En la Figura 3.3 se comprueba que los ficheros de datos contienen imágenes (cuadro amarillo de la Figura 3.3) e información útil para futuras líneas de investigación. En este trabajo nos limitaremos a tomar la información correspondiente al número de escaños por partido político. Añadido está el problema de las coaliciones, por ejemplo, Podemos está separado en: PODEMOS-IU-EQUO, PODEMOS-COM, PODEMOS-EN MAREA-ANOVA-EU, ... lo cual complica más todavía el uso de los datos directamente desde el fichero excel.

Para solventar el problema de las coaliciones de Podemos, problema que sólo aparece en los dos últimos ficheros ya que sólo han estado presentes en las elecciones de Junio 2016 y Diciembre 2015, se ha recurrido a [18] para agrupar diferentes opciones políticas. La idea de agrupar participantes ha sido debido a la complejidad computacional de los algoritmos para calcular los índices de poder. A medida que aumenta el número de participantes en las votaciones el tiempo crece exponencialmente, por lo que se ha considerado que era más representativo juntar las coaliciones que excluir a partidos con pocos escaños, como podría ser Coalición Canaria. La ejecución, por ejemplo en 2016, con 9 partidos calcula $9! = 362{,}880$ permutaciones y $2^8 \times 9 = 2304$ combinaciones, lo cual invita a pensar que el algoritmo tardará bastante tiempo en ejecutarse.

Para resolver el problema de los ficheros corruptos y poco estructurados, se tomó la decisión de modificar manualmente los archivos excel descargados de la página del Ministerio [2]. Dado que suponía modificarlos y leerlos con una librería de R, se decidió introducir a mano los datos.

3.3. Algoritmos

En esta sección se analizarán los algoritmos desarrollados a lo largo de este trabajo para la aplicación del cálculo de los índices de poder.

3.3.1. Algoritmo desarrollado para el cálculo de índices de poder en juegos (3,2)

En esta sección se va a abordar el problema de la combinación de juegos (2,2) para conseguir un juego (3,2). Se recuerda que un juego (3,2) es un SVP en el que las opciones de voto son sí, no y abstención; y un juego (2,2) es un SVP con tan sólo sí y no como opciones de voto. En ambos casos la medida puede ser aprobada o rechazada. Este algoritmo es contribución de este TFG.

Definición **3.1.** *Sea J un Juego (3,2) con n participantes configurado con mayoría simple.*

Llamaremos R_i *al Juego (2,2) con* $n-1$ *participantes configurado con mayoría absoluta en el que el jugador* $i-$*ésimo no participa (se considera que es el participante que se abstiene).*

Definición **3.2.** *Sea J un Juego (3,2) con n participantes,* $x_1, ..., x_n$*, configurado con mayoría simple. Sea* $I = \{i_1, i_2, ..., i_k\}$ *un conjunto de índices. Se llamará, de forma análoga a la definición 3.1,* $R_{i_1, i_2, ... i_k}$ *al juego (2,2) con* $n-|I|$ *participantes configurado con mayoría absoluta, en el que el jugador* x_i *no participa si* $i \in I$.

Proposición 2. *Sea J un Juego (3,2) con n participantes,* $x_1, ..., x_n$*, configurado con mayoría simple.*

Supongamos que en la primera votación x_1 *se abstiene, por lo que tenemos un juego (2,2) con mayoría absoluta de* $n-1$ *participantes,* R_1*. Se analiza el poder en este caso y se acumula el poder de cada participante* $x_2, ..., x_n$.

Se repite el proceso con los $n-1$ *participantes siguientes,* $x_2, ...x_n$*, teniendo* R_i *para cada* x_i *absteniéndose.*

A continuación, se sigue el proceso absteniendo de dos en dos participantes: primero x_1 *y* x_2*, después* x_1 *y* x_3*, ..., sin repeticiones. Obteniendo* $R_{1,2}$*,* $R_{1,3}$*, ... Se analiza cada caso y se acumula el poder de cada participante al ya calculado en pasos previos.*

Se continúa con la misma metodología, hasta abstener $n-1$ *participantes, pues abstener a todos no tendría sentido.*

Para dar una fórmula, es preciso definir $C_{i,j}$ *como el conjunto de combinaciones de i elementos escogidos de j en j, sin repeticiones. Estas combinaciones se representarán de la siguiente forma: 1 2, 1 3, ..., 2 3, 2 4, ..., 3 4, ..., si j fuera 2.*

Ejemplo **3.1.** *Veamos un ejemplo de* $C_{i,j}$ *particular:*

$$\text{Sea } X = \{1,2,3,4\}. \text{ Entonces } C_{4,2} = \begin{cases} 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 2 & 3 \\ 2 & 4 \\ 3 & 4 \end{cases}$$

Matemáticamente, $C_{i,j}$ *representa un número,* $\binom{i}{j}$*, pero en este trabajo, por conveniencia, se denota* $C_{i,j}$ *como el conjunto de las combinaciones, no como dicho número. Es decir, es un conjunto de combinaciones ordenadas sin repeticiones. Nótese que* $|C_{i,j}| = \binom{i}{j}$.

Se recuerda que J es el juego (3,2). Se muestra como quedaría la combinación de éste como juegos (2,2).

$$J = \sum_{i=0}^{n-1} \sum_{\substack{j \in I \\ I = C_{n,i}}} R_j \tag{3.1}$$

Hay que señalar que en la ecuación 3.1 los índices j del segundo sumatorio son combinaciones, es decir, $j = 1$ ó $j = 1\ 2\ 3$, pero en ambos casos $R_{i_1\ i_2\ i3}$ se interpreta como en la definición 3.2.

Para entender un poco mejor esta ecuación se procede a analizar primero el sumatorio más interno. Se muestran los valores que toma I.

- Si $i = 0$, $I = \emptyset$, caso particular donde todos votan sí/no.
- Si $i = 1$, $I = \{1, ..., n\}$
- Si $i = 2$, $I = \{1\ 2, 1\ 3, ..., 1\ n, 2\ 3, 2\ 4, ..., 2\ n, ...\}$
- Si $i = 3$, $I = \{1\,2\,3, 1\,2\,4, ..., 1\,2\,n, 1\,3\,4, 1\,3\,5, ..., 1\,3\,n, 2\,3\,4, 2\,3\,5, ..., 2\,3\,n, 2\,4\,5, 2\,4\,6, ..., 2\,4\,n, ...\}$
- ...
- Si $i = n-1$, $I = \{1\ 2\ 3\ ...\ n-1, 1\ 2\ 3\ ...\ n, 1\ 3\ 4\ ...\ n, 1\ 2\ 4\ ...\ n, 1\ 2\ 3\ 5\ ...\ n, 2\ 3\ 4\ ...\ n\}$

Por simplificar más, se exponen los sumatorios representativos, sin límites del sumatorio. Simplemente reflejando los grupos de abstención R_i, $R_{i,j}$, ...

$$J = R_0 + \sum_i R_i + \sum_{i,j} R_{i,j} + \sum_{i,j,k} R_{i,j,k} + ... + \sum_{i,j,k,...,l} R_{i,j,k...,l} \tag{3.2}$$

- El primer sumatorio en un juego (2,2) de mayoría absoluta con n participantes.
- El segundo sumatorio son $\binom{n}{1} = n$ juegos (2,2) de mayoría absoluta a evaluar. Cada juego con $n-1$ participantes.
- ...
- El $i-\acute{e}simo$, suponiendo $i \leq \frac{n}{2}$, sumatorio son $\binom{n}{i} = \frac{n \times (n-1) \times ... \times (n-i+1)}{i!}$ juegos (2,2) de mayoría absoluta a evaluar. Cada juego con $n-i$ participantes.
- ...
- El $i-\acute{e}simo$, suponiendo $i > \frac{n}{2}$, sumatorio son $\binom{n}{i} = \frac{n \times (n-1) \times ... \times (i+1)}{(n-i)!}$ juegos (2,2) de mayoría absoluta a evaluar. Cada juego con $n-i$ participantes.
- ...
- El último sumatorio son $\binom{n}{n-1} = n$ juegos (2,2) de mayoría absoluta a evaluar. Cada juego con 1 participante.

Ahora bien, esta equivalencia entre un juego (3,2) de mayoría simple y varios juegos (2,2) de mayoría absoluta no ha solucionado el coste computacional, pues hay que realizar muchas combinaciones de juegos (2,2) de mayoría absoluta. A pesar de que en los juegos (2,2) de mayoría absoluta se calculan los índices de poder más rápido pues tienen menos opciones de voto, el número tan elevado de juegos hace que la cantidad de cálculos sea demasiado costoso. Por todo esto, se ha decidido utilizar la proposición 1 en lugar de este algoritmo, ya que permite un ahorro mayor de tiempo computacional.

3.3.2. Algoritmo de Borda

Este índice, como se ha expuesto en la sección 2.2.2 es el más simple. La clave de este algoritmo es decidir el peso que se le dará a cada valor obtenido en cada vuelta de las elecciones. Como se indicó en el apartado 2.2.2, se ha dado peso: 10, 5 y -1 a la primera, segunda y tercera vuelta, respectivamente.

Por tanto, el algoritmo consiste en multiplicar los valores obtenidos en cada una de las vueltas por el peso que se le da a cada una de éstas. El valor obtenido es el poder que tiene el participante.

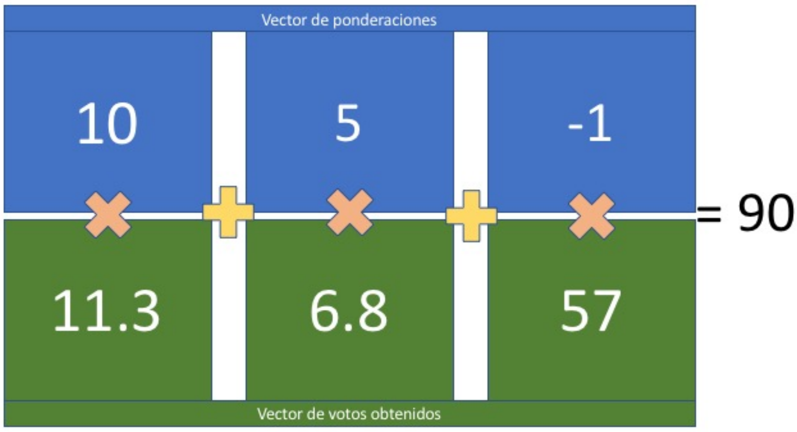

Figura 3.4: Esquema del cálculo del índice de Borda para un partido

En la Figura 3.4 vemos un ejemplo del algoritmo: En la primera vuelta el partido x ha conseguido 11,3 votos, en la segunda ha logrado 6,8 y en la tercera 57. Dado que los pesos son 10, 5 y -1, se obtiene el valor de: $10 \times 11,3 + 5 \times 6,8 + (-1) \times 57 = 90$.

Puesto que este algoritmo es el más simple y el que menos refleja el sistema de votación español sólo se ha realizado para los cuatro partidos con mayor número de escaños: PP, PSOE, Podemos y Ciudadanos; y para las encuestas preelectorales al Congreso de los Diputados de 2016.

3.3.3. Algoritmo de Banzhaf

Este algoritmo ha pasado por diferentes fases en todo el desarrollo de este trabajo. Primero empezó como un algoritmo estático, que calculaba el IB para 3 ó 4 participantes. Pronto se quedó corto, por lo que se amplió a un algoritmo que calculase el IB para cualquier número de participantes. Por ello, se observó la necesidad de tener ordenados de mayor a menor el vector de los votos obtenidos. De dicho modo se ahorraban muchas combinaciones. Además, es una hipótesis lógica y es más intuitivo tener $(10, 5, 4, 2, 1)$. Esta observación tiene una simple explicación desarrollada en este trabajo:

Demostración. Sea J un juego (2,3) con cuota q y n participantes $x_1, ...x_n$, representados por sus votos, es decir, cada x_i es el valor $i - \acute{e}simo$ del vector de votos obtenidos.

Supongamos que tenemos una coalición $A = \bigcup_{i=1}^{l} x_i$, con x_{i_0} votante basculante en dicha votación. Supongamos que el poder de x_{i_0} por ahora es p, siendo p el poder acumulado en las evaluaciones previas, sin normalizar.

x_1	x_2	...	x_{i_0}	...	x_l

Tabla 3.1: Combinación de participantes x_i

Vamos a añadir a la coalición A el participante x_{l+1}. Denotaremos $\tilde{A}$ a esa nueva coalición, quedando $\tilde{A} = A \cup x_{l+1}$.

Si x_{i_0} sigue siendo basculante en esta nueva votación, significa que:

$$\sum_{j=0,\ j\neq i_0}^{l+1} x_j = \sum_{j=0,\ j\neq i_0}^{l} x_j + x_{l+1} < q \tag{3.3}$$

Como el vector de votos está ordenado de manera descendente, tenemos $x_i \leq x_j, \forall i \geq j$. Luego, si el término x_{l+1} es sustituido por otro x_{l+s}, con $n \geq s > 1$, seguimos teniendo que la suma de los votos es menor a la cuota, pues si en la ecuación 3.3 sustituimos x_{l+1} por x_{l+s} tenemos:

$$q > \sum_{j=0,\ j\neq i_0}^{l} x_j + x_{l+1} > \sum_{j=0,\ j\neq i_0}^{l} x_j + x_{l+s}\ \forall n \geq s > 1\ , \tag{3.4}$$

por lo que $\tilde{A} \setminus x_{i_0}$ no sería una coalición ganadora. Pero teníamos que $\tilde{A}$ si era ganadora. Por tanto, x_{i_0} sigue siendo basculante en todas las combinaciones de un elemento con índice mayor que l. Esta estrategia ahorra evaluar $n-l$ votaciones particulares, pues sabemos que el poder de x_{i_0} se incrementará en $\binom{n-l}{1}$ sin necesidad de analizar cada una de las combinaciones. Del mismo modo se sigue con $\tilde{A} = A \cup x_{l+1} \cup x_{l+2}$, pues si x_{i_0} vuelve a ser basculante, sabremos que su poder se incrementará en $\binom{n-l-2}{2}$.

La fórmula final para la coalición $\tilde{A}$ nos quedaría:

$$\tilde{A} = A \bigcup_{k=1}^{n-l} x_{l+k}\ , \tag{3.5}$$

y el incremento de poder para el participante x_{i_0} quedaría:

$$p = p + \binom{n-l}{k}\ , \tag{3.6}$$

siendo k el número de participantes que añadimos a $\tilde{A}$

El análisis sería válido para cualquier x_i.

□

El algoritmo de Banzhaf descrito hasta ahora en este apartado, sólo contempla juegos (2,2) configurados con mayoría absoluta. Por lo que se decidió seguir investigando hacia un algoritmo que tuviera en cuenta la mayoría simple y dejar que los participantes pudiesen votar sí/no/abstención. Para esta nueva situación no es válido el algoritmo anterior, pues simplemente agrupa participantes sumando sus votos, suponiendo que el resto de participantes votan no.

El desarrollo de este algoritmo empezó en una hoja Excel, con un juego (3,2) con pocos participantes. Lo primero observado fue que una codificación sencilla de los votos simplificaba considerablemente los cálculos. Más concretamente, el voto positivo de un partido se codificaría como 1, el voto negativo como -1, y la abstención como 0. Por tanto, multiplicando cada una de las codificaciones por el número de votos de cada partido se obtiene un número entero que, en caso de ser positivo, la medida es aprobada y, en caso negativo o nulo, la medida es rechazada. Notar que con esta aproximación no es necesario obtener un valor explícito para la cuota. Esto es debido a que en mayoría simple, para aprobar una medida sólo es necesario tener mayor números de votos del tipo sí que del tipo no. Por tanto, si multiplicamos todos los sí por 1, no por -1 y abstención por 0, el signo de la suma indicará qué tipo de voto ha sido mayoritario. Si es positivo ganan los sí, en cambio, si es negativo ganan los noes.

Por tanto, para realizar el algoritmo es necesario generar todas las combinaciones de n elementos, siendo n el número de participantes de un juego (3,2), con valores $-1, 0, 1$ e ir evaluando participante a participante si cambia el voto entre sí y no, si esto implica un cambio de estado de ley, es decir, si pasa de aprobarse a bloquearse, o viceversa. Se muestra el procedimiento con un ejemplo seleccionando a los cuatro partidos con mayor número de escaños:

Ejemplo **3.2.** *Supongamos que tenemos el siguiente vector de votos: (137, 85, 71, 32) y configuración de mayoría simple. Para ilustrar el procedimiento, generamos tres posibles combinaciones:*

Combinación	PP	PSOE	Podemos	Ciudadanos	Decisión	Signo
1	1	1	1	1	325	+
2	-1	-1	-1	1	-261	-
3	1	-1	0	0	52	+

Tabla 3.2: Tabla del ejemplo 3.2

Tendríamos que realizar el proceso que vamos a detallar a continuación para cada uno de los participantes, pero en este caso lo haremos sólo para el primero, que corresponde con el PP.

Vamos a evaluar cada uno de los tres casos particulares que se muestran en la Tabla 3.2:

Combinación 1. *Si cambiamos de bando al primer participante queda un vector (-1, 1, 1, 1). Y la decisión tendría un valor de 51, -1* $\times 137 + 1 \times 85 + 1 \times 71 + 1 \times 32 = 51$.

Como el signo de la decisión es el mismo, positivo, la ley no cambia de estado, pues para ambas opciones del participante la ley se mantiene, luego el participante no es basculante en esta votación.

Combinación 2. *Si cambiamos de bando al primer participante queda el vector (1, -1, -1, 1). Y la decisión tendría un valor de 13, 1* $\times 137 + (-1) \times 85 + (-1) \times 71 + 1 \times 32 = 13$.

Como el signo de la decisión es distinto, la ley cambia de estado. Pasa de ser rechazada a ser aprobada, luego el participante es basculante en esta votación. Su poder se incrementa en uno.

Combinación 3. *Si cambiamos de bando al primer participante nos queda el vector (-1, -1, 0, 0). Y la decisión tendría un valor de -222, -1* $\times 137 + (-1) \times 85 + 0 \times 71 + 0 \times 32 = -222$.

Como el signo de la decisión es distinto, la ley cambia de estado. Pasa de ser aprobada a ser rechazada, luego el participante es basculante en esta votación. Su poder se incrementa en uno.

Con estas evaluaciones hemos calculado el índice de poder en estas tres combinaciones del primer participante. Tendríamos que realizar el mismo proceso con cada uno de los participantes restantes, y con el resto de combinaciones posibles.

No hace falta considerar aquellos casos en los que el participante evaluado vota abstención, puesto que tal y como se vio en la proposición 1 (apartado 2.3.2) es equivalente el IB evaluando al participante $i - \acute{e}simo$ con las tres opciones de voto como con sólo sí/no.

Otra decisión importante que se tomó fue el estado que adquiriría una medida si ante una votación todos los participantes se abstienen. El valor de la votación sería 0, que no tiene signo. Entonces, ¿se aprueba o se deniega la propuesta?. Se decidió que se denegaría la propuesta porque por defecto la ley no está aprobada y sólo se aprobaría si se consiguen reunir más votos a favor que en contra, por lo que ningún voto a favor implica la no aprobación de la medida. En realidad esta decisión sólo tiene sentido si se evalúan los cambios cortos y largos del participante $i - \acute{e}simo$ porque si se tiene en cuenta la proposición 1, los cambios cortos no serían necesarios, por lo que ese caso base nunca se daría.

3.3.4. Algoritmo de Shapley-Shubik

Este es el tercer y último algoritmo que se analizará. Es más complicado que el IBR pero un poco más sencillo que el IB. Para este algoritmo se vuelve a una configuración de los juegos (2,2) en mayoría absoluta, por lo que no se permiten las abstenciones.

Para su implementación se han generado todas las permutaciones posibles del vector de votos obtenidos de los participantes en el juego (2, 2). Esto ha sido sencillo, pues el lenguaje R, implementa funciones que se encargan de ello. Pero entender dichas funciones, y los tipos de datos con los que trata ha sido bastante más complicado.

Una vez se han generado las permutaciones el algoritmo es simple: se tiene que evaluar individualmente cada una de las permutaciones, decidiendo en cada caso qué posición es la pivotante (posición que suma los votos necesarios para cambiar el estado de la ley).

Se muestra su funcionamiento en el siguiente ejemplo:

Ejemplo **3.3.** *Supongamos, igual que en el ejemplo 3.2, que tenemos el vector de votos: (137, 85, 71, 32).*

La cuota de este juego (2,2) sería de $\lfloor \frac{137+85+71+32}{2} \rfloor + 1 = 163$.

Como tenemos cuatro participantes, sabemos que el número de casos posibles de permutaciones son $4! = 4 \times 3 \times 2 \times 1 = 24$ *casos. Luego, el valor de normalización será* 24.

Permutación	e_1	e_2	e_3	e_4
1	137	85	71	32
2	32	85	137	71
3	71	85	32	137
4	137	71	32	85

Tabla 3.3: Tabla de permutaciones del ejemplo 3.3

En la tabla 3.3 hemos generado cuatro permutaciones de ejemplo y vamos a analizar cada una de ellas.

Permutación 1. *En esta primera permutación, el elemento que permite superar, o igualar, la cuota, 163, es el segundo. Pues* $137 < 163$ *y* $137 + 85 = 222 > 163$. *Luego, el participante que está en la posición dos de la permutación 1 suma el valor de uno en su acumulador de poder (este participante es el PSOE, pues es el votante que cuenta con 85 votos).*

Permutación 2. *En la segunda permutación, el elemento que permite superar, o igualar, la cuota, 163, es el tercero. Pues* $32 < 163$, $32 + 85 = 117 < 163$ *y* $32 + 85 + 137 = 254 > 163$. *Luego, sumamos uno en el acumulador de poder del participante que ocupa la posición tercera en la permutación 2 (este participante es el PP, que es el participante con 137 votos).*

Permutación 3. *En la tercera permutación, el elemento que permite superar, o igualar, la cuota, 163, es el tercero también. Pues* $71 < 163$, $71+85 = 156 < 163$ *y* $71+85+32 = 188 > 163$. *Luego, sumamos uno en el acumulador de poder del participante que ocupa la posición tercera en la permutación 3 (este participante es Ciudadanos, que es el participante con 32 votos).*

Permutación 4. *En la cuarta permutación, el elemento que permite superar, o igualar, la cuota, 163, es el segundo. Pues* $137 < 163$ *y* $137 + 71 = 208 > 163$ *. Luego, sumamos uno en el acumulador de poder del participante que ocupa la posición segunda en la permutación 2 (este participante es Podemos, que es el participante con 71 votos).*

Finalmente, tenemos que el poder queda de la siguiente manera:

PP	**PSOE**	**Podemos**	**Ciudadanos**
4 %	4 %	4 %	4 %

Tabla 3.4: Tabla de valores del índice de poder de Shapley-Shubik

La finalidad de este ejemplo no es más que mostrar el funcionamiento del algoritmo, ya que los valores obtenidos no son representativos, pues necesitaríamos hacer las 24 permutaciones.

4

Análisis resultados

En este capítulo se van a presentar los resultados obtenidos tras analizar los datos recopilados de Twitter así como las gráficas generadas por la aplicación en Shiny para cada uno de los tres índices de poder.

4.1. Probabilidades de coalición

En esta sección se mostrarán los resultados obtenidos tras analizar los datos de Twitter. Estos fueron descargados a fecha 24 de septiembre de 2016.

Primeramente se muestran algunos datos característicos:

	Partidos políticos	**Políticos**
Número de usuarios	2.855.928	3.879.139

Tabla 4.1: Total de usuarios de Twitter.

Como se aprecia en la Tabla 4.1, hay del orden de más de 2 millones de cuentas que siguen a partidos políticos, y más de 3 millones a representantes, lo cual ya refleja que hay aproximadamente un 35 % más de cuentas que siguen a representantes de los partidos políticos que a los partidos.

En la Tabla 4.2 se muestran el total de usuarios que siguen a cada político, y el porcentaje que representa cada uno:

Político	**Usuarios**	
Pablo Iglesias	1.86 millones	47.93 %
Mariano Rajoy	1.27 millones	32.73 %
Pedro Sanchez	0.39 millones	10.27 %
Albert Rivera	0.35 millones	9.06 %

Tabla 4.2: Datos de cuentas que siguen a políticos.

Claramente Pablo Iglesias (PI) es el usuario por excelencia en cuanto a seguidores, pues acumula alrededor del 50 % de los usuarios de Twitter que siguen a uno de estos cuatro políticos.

Estos datos no significan que PI sea el representante más querido, sino el más seguido. Conviene destacar que seguir a un usuario en Twitter no implica que éste tenga una ideología afín a la propia. En Twitter es complicado obtener un patrón que modelice las razones de hacer *follow* y *unfollow* (seguir y dejar de seguir) a cierto usuario X.

A continuación, en la Tabla 4.3 se mostrarán los datos referidos a los partidos políticos:

Partido político	**Usuarios**	
Podemos	1.14 millones	39.79 %
PP	0.58 millones	20.22 %
Ciudadanos	0.68 millones	23.73 %
PSOE	0.46 millones	16.26 %

Tabla 4.3: Datos de cuentas que siguen a partidos políticos.

En la Tabla 4.3 se atenúan las diferencias observadas entre los partidos políticos en cuanto al número de seguidores de sus lideres. Podemos es el partido dominante.

En este primer análisis se observa la limitación ya mencionada en la sección 2.1.1 del sesgo de Twitter. Tanto Pablo Iglesias, en el ámbito de políticos, como Podemos en el contexto de partidos, son los usuarios por excelencia, teniendo casi el 50 % y el 40 % respectivamente, mostrando el sesgo que se menciona en el artículo [8]. Este artículo muestra como los usuarios de Twitter son en mayor parte seguidores de Podemos, seguido por Ciudadanos.

A continuación, las Tablas 4.4, 4.5 y 4.6 muestran las tres matrices A, B y C, respectivamente, descritas en la sección 2.1.2 y calculadas a partir de los datos obtenidos mediante la herramienta desarrollada en Python. Se estructurarán en forma de tabla para que sean más comprensibles.

	P(PP | FILA)	P(PSOE | FILA)	P(P | FILA)	P(CS |FILA)
Dado que sigue al PP	1	0.466	0.489	0.329
Dado que sigue al PSOE	0.580	1	0.573	0.365
Dado que sigue a P	0.248	0.234	1	0.169
Dado que sigue a Cs	0.281	0.250	0.284	1

Tabla 4.4: Matriz de probabilidades de partidos condicionadas a partidos.

	P(MR | FILA)	P(PS | FILA)	P(PI | FILA)	P(AR |FILA)
Dado que sigue al PP	0.715	0.251	0.561	0.240
Dado que sigue al PSOE	0.579	0.439	0.630	0.274
Dado que sigue a P	0.323	0.162	0.842	0.124
Dado que sigue a Cs	0.424	0.335	0.461	0.403

Tabla 4.5: Matriz de probabilidades de políticos condicionadas a partidos.

	P(PP | FILA)	P(PSOE | FILA)	P(P | FILA)	P(CS |FILA)
Dado que sigue a MR	0.325	0.212	0.289	0.226
Dado que sigue a PS	0.364	0.512	0.463	0.570
Dado que sigue a PI	0.152	0.143	0.515	0.103
Dado que sigue a AR	0.395	0.362	0.401	0.778

Tabla 4.6: Matriz de probabilidades de partidos condicionadas a políticos.

En estas tres matrices, A, B y C, se puede apreciar que las probabilidades de pacto entre partidos no pueden ser estimadas a partir de estas probabilidades condicionadas dado que, por ejemplo, en la matriz A (Tabla 4.4) se refleja que la probabilidad de pacto entre PP y Podemos es muy alta, un 48,9 %, mientras que la realidad parece indicar todo lo contrario. También es destacable, en esa misma matriz, el valor obtenido para la probabilidad de pacto entre PSOE y Podemos, pues es similar a la probabilidad de pacto entre PP y Podemos, lo cual apoya la hipótesis de la poca representatividad de estas métricas. Un argumento más en contra de utilizar este sistema. Por tanto, finalmente se decidió considerar la hipótesis de equiprobabilidad de coalición entre todos los partidos para evaluar los algoritmos descritos en la sección 3.3.

Estos datos presentan más conclusiones destacables, por ejemplo, que la gente que sigue al PSOE sigue más a Pablo Iglesias que a su propio representante (Tabla 4.5), lo cual es digno de mencionar. También destaca que Albert Rivera es claramente el líder que más atrae a los propios seguidores de su partido, obteniéndose que el 77,80 % de los seguidores de la cuenta de Ciudadanos sigue también a Albert Rivera (Tabla 4.6). Es curioso comprobar como los usuarios que siguen a Mariano Rajoy siguen en igual medida a cada uno de los tres partidos políticos restantes.

4.2. Índices de poder

En esta sección se mostrarán las gráficas obtenidas con cada uno de los índices de poder estudiados con los datos de las elecciones al Congreso de los Diputados. Estas gráficas serán histogramas pues se ha considerado que eran la representación gráfica más acertada dada la estructura de los datos. Añadido a que visualmente es muy revelador, ya que viendo las barras, sin ver el valor real, el usuario se puede hacer una idea de la distribución del poder.

A continuación se mostrará una gráfica por apartado, salvo en IB, que tendrá dos, una por cada configuración posible. En el caso de IB e ISS se han seleccionado los datos correspondientes al último año de las elecciones al Congreso de los Diputados, es decir, 2016 porque son las más actuales.

4.2.1. Índice de Borda

En este caso sólo se disponen de datos de una encuesta del CIS [1], por lo que la aplicación sólo genera la gráfica que se muestra a continuación.

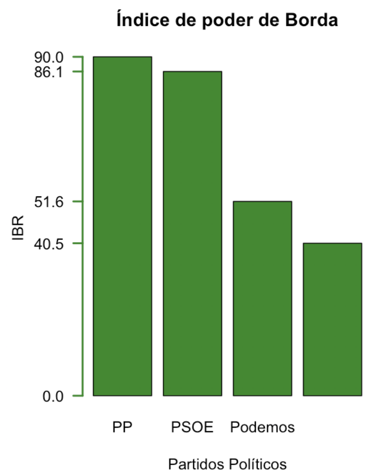

Figura 4.1: Histograma del índice de Borda

Como se ve en la Figura 4.1 el partido con mayor poder es el PP, seguido por el PSOE a tan solo 3,9 puntos de diferencia. Podemos y Ciudadanos tienen bastante menos poder. Estos resultados coinciden con el panorama actual de política, ya que el Presidente de Gobierno corresponde con el partido con mayor poder. Sin embargo, la casi igualdad en cuanto a escaños entre PSOE y Podemos no se refleja en esta gráfica, ya que PSOE y Podemos tienen una diferencia de poder de 34,5.

4.2.2. Índice de Banzhaf

Para este índice se han generado 26 gráficas correspondientes a las 13 elecciones al Congreso de los Diputados, dos por cada elección: una permitiendo votar abstención y otra sin permitirlo. Sólo se mostrará la correspondiente a las elecciones al Congreso de los Diputados de 2016 al ser las elecciones más cercanas. Con el fin de mejorar la visualización, se han utilizado las siguientes abreviaturas para algunos de los partidos políticos.

- C's corresponde a Ciudadanos.
- ERC corresponde a Esquerra Republicana de Catalunya.
- CDC corresponde a Convergencia Democrática de Cataluña.
- EAJ-PNV corresponde a El Partido Nacionalista Vasco.
- CCa corresponde a Coalición Canaria.

Estas abreviaturas serán utilizadas para las Figuras 4.2 y 4.3 mostradas en esta sección, como para la Figura 4.4 que se presenta en la sección 4.2.3.

Se procede a mostrar cada una de las gráficas obtenidas, siendo la primera la correspondiente a las elecciones sin abstención y la segunda permitiendo a los votantes la opción de voto en abstención.

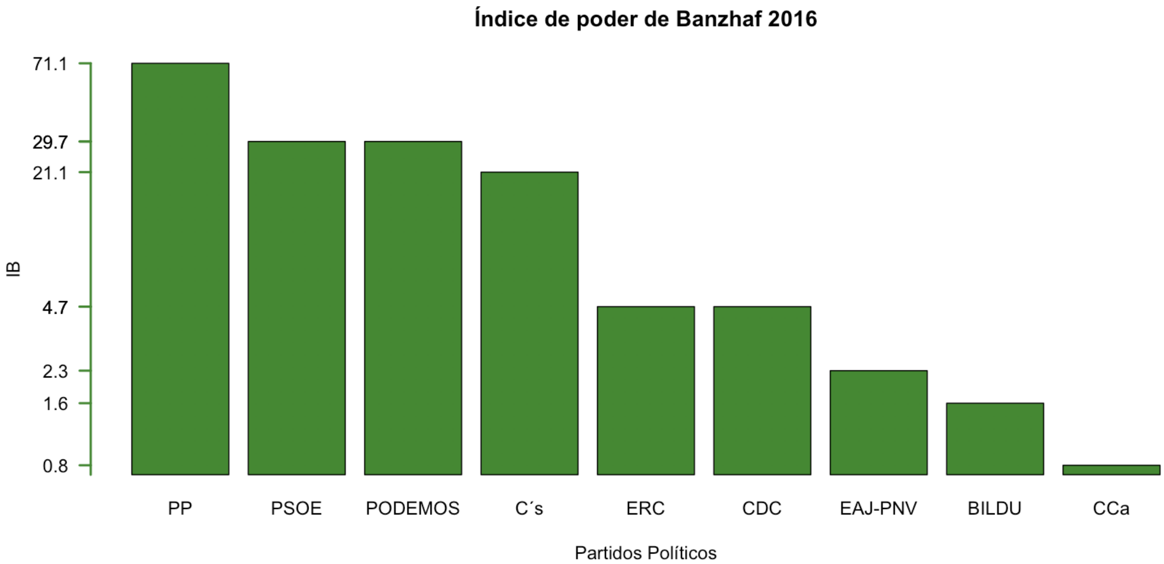

Figura 4.2: Histograma del índice de Banzhaf sin abstención

La Figura 4.2 representa el índice de poder de Banzhaf de las elecciones en 2016 configurados con mayoría absoluta y sin contemplar abstenciones. Los valores en el eje y (IB) son porcentajes, normalizando por 2^n, y la escala escogida es la logarítmica, pues al tener valores tan pequeños los resultados en escala lineal serían difícilmente analizables. Se aprecia que el PP es el más poderoso y que el PSOE y Podemos tienen el mismo poder a pesar de tener 85 y 71 votos, respectivamente. Es destacable que tengan el mismo poder a pesar de tener una diferencia de 14 votos. ERC y CDC también tienen el mismo poder, pero en este caso cuentan con 9 y 8 votos, por lo que parece más natural que tengan mismo poder pues sólo tienen un voto de diferencia. Podría incluso decirse que Coalición Canaria (CCa) es prácticamente un hombre de paja, pues su poder es de 0,8 %, lo que significa que en las decisiones no tiene apenas poder de decisión.

A continuación, en la Figura 4.3 se representará el poder de Banzhaf de las elecciones al Congreso de 2016 configuradas con mayoría simple. Este ejemplo representa la segunda ronda de votación en la sesión de investidura. Los valores en el eje y (IB) son porcentajes, normalizados por 3^n, y la escala escogida es la logarítmica por la misma razón que en la Figura 4.2. Los acrónimos de los partidos están definidos en la lista redactada anteriormente.

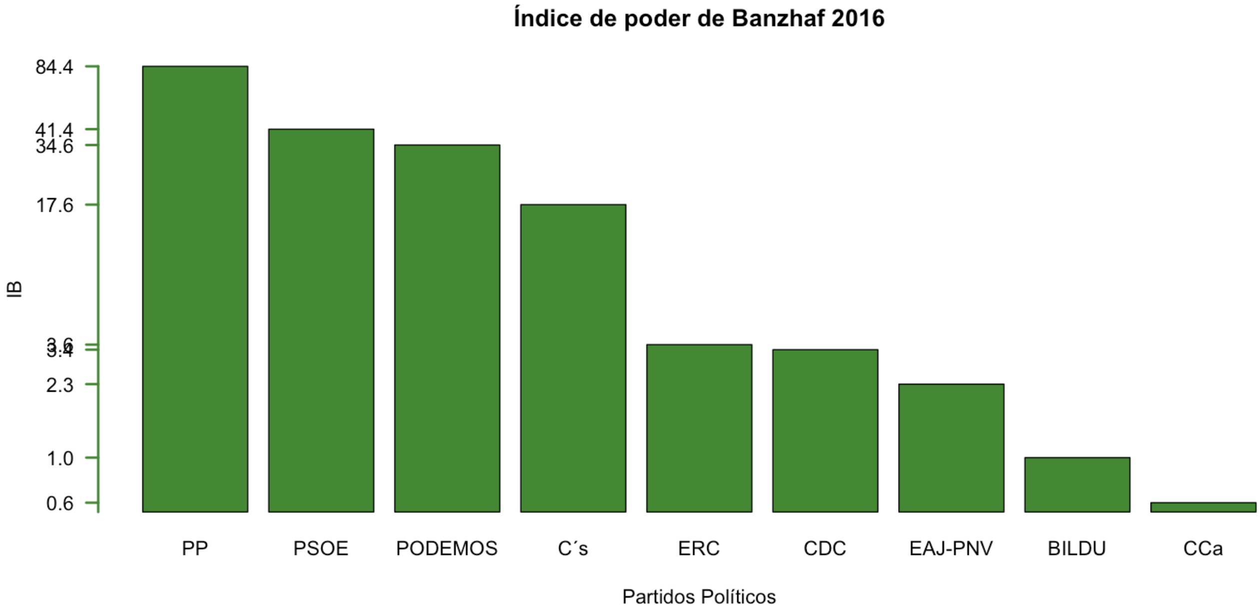

Figura 4.3: Histograma del índice de Banzhaf con abstención

El partido con mayor poder sigue siendo el PP, ganando incluso poder (gana un 13 %); mientras que el PSOE y Podemos ya no tienen el mismo poder, puesto que PSOE gana un 11 % y Podemos tan solo gana un 5 %. Si antes Coalición Canaria se podía considerar cercana a ser hombre de paja, permitiendo abstenciones esta consideración es más cierta, pues pierde un 0,2 % de poder, quedándose con un 0,6 %.

Con esta gráfica también se puede destacar que una votación en la que la abstención es una opción de voto, PP es el mayor beneficiado, al igual que ocurrió en las investidura al Presidente de Gobierno. Sin embargo, Ciudadanos ha perdido poder pues sin abstenciones tenía un 21,1 % mientras que en este juego con abstenciones ha obtenido un 17,6 %, es decir, ha perdido 3,5 puntos de poder. Por último, se aprecia que ERC y CDC ya no tienen exactamente el mismo poder, pero si aproximadamente, variando unas décimas.

4.2.3. Índice de Shapley-Shubik

Para este índice se han generado 13 gráficas, correspondientes a las 13 elecciones al Congreso de los Diputados, pero sólo se mostrará la correspondiente a las elecciones al Congreso de los Diputados de 2016 dada su cercanía y la complejidad del escenario político que se planteó para la elección del Presidente del Gobierno.

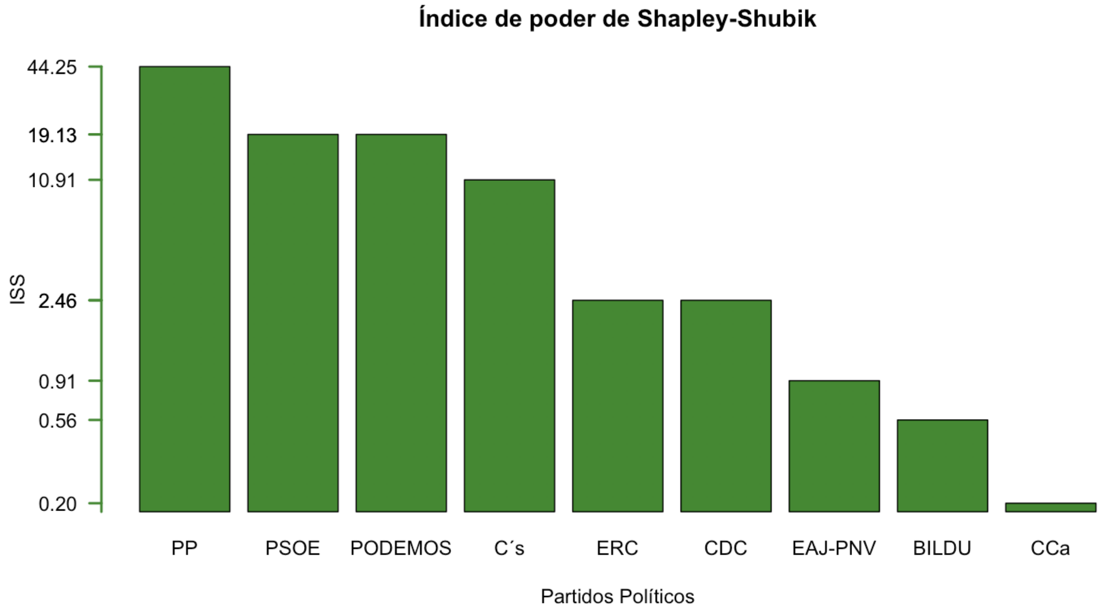

Figura 4.4: Histograma del índice de Shapley-Shubik

En la Figura 4.4 se aprecia un resultado similar al de la Figura 4.2. Los valores en el eje y (IS) se entienden como porcentajes, normalizados entre la suma total de los valores de los índices de poder, y la escala escogida es la logarítmica. Los acrónimos están definidos en la lista del apartado 4.2.2.

PP sigue siendo el partido político con mayor poder, alcanzando un 44,25 %, que es inferior al poder alcanzado con el índice de poder de Banzhaf que era 71,1 % sin abstención y 84,4 % con abstención. El índice de poder de PSOE y Podemos está más cercano al de PP que en la sección 4.2.2. Según esta gráfica se puede concluir que el poder está repartido entre: PP, PSOE, Podemos y Ciudadanos, teniendo en lugar de bipartidismo algo cercano a pluripartidismo.

5

Conclusiones y trabajo futuro

En este Trabajo de Fin de Grado se han diseñado dos herramientas que permiten calcular probabilidades de coalición e índices de poder. La aplicación de estas herramientas a datos reales ha permitido concluir argumentos bastante interesantes que se exponen a continuación.

5.1. Conclusiones

5.1.1. Probabilidades de coalición

Primeramente se obtuvieron conclusiones respecto al análisis de datos de Twitter para estimar probabilidades de coalición, como se ha desarrollado en la sección 4.1. La conclusión obtenida es que esta teoría desarrollada no permite estimar probabilidades de coalición porque los datos obtenidos no representan el panorama de la política actual, ya que una de las probabilidades de coalición más alta es obtenida en la alianza PP-Podemos, algo imposible en la práctica ya que son de ideales opuestos.

5.1.2. Herramienta de cálculo de índices de poder

La herramienta desarrollada en Shiny ha permitido realizar el cálculo de los tres índices de poder desarrollados en este trabajo. En particular, se ha configurado con datos de las elecciones al Congreso de los Diputados en España desde 1977 hasta 2016, pero también permite la introducción manual de datos según quiera el usuario.

5.1.3. Propuesta de algoritmos para el cálculo eficiente de índices de poder en juegos (3,2)

Se ha realizado un estudio de los juegos (3,2) permitiendo desglosarlos en juegos (2,2) donde no se tiene en cuenta la abstención. Tras este análisis se ha concluido que esta descomposición no ahorra suficiente tiempo computacional, sólo aporta otra visión sobre los juegos (3,2), ya que no evita tener que calcular un número muy elevado de juegos (2,2).

Finalmente, se ha concluido que en los juegos (3,2) se puede evaluar a cada participante para calcular su poder cambiando su opción de voto entre sí y no, sin tener en cuenta para ese votante la opción de abstención, según lo visto en la proposición 1 en la sección 2.3.2.

5.1.4. Análisis y estudio de índices de poder

Con respecto a los índices de poder, se obtienen resultados que representan aproximadamente el panorama actual de la política, pues PP es el partido con mayor poder seguido por PSOE y Podemos, que incluso en algunos casos empatan en poder; y el cuarto en este ranking de índices de poder sería Ciudadanos. El índice que menos representa el panorama actual es el índice de Borda. Quizá esto sea debido a los pesos escogidos y a su simpleza. Al contrario, los índices de Banzhaf y Shapley-Shubik no sólo representan el panorama actual, sino que son muy similares entre ellos y aportan información semejante. También se ha podido concluir que en las votaciones con la opción de abstención se incrementa el índice de poder de algunos participantes, por ejemplo el PP vió incrementado su poder en un 13 %. Por lo que es interesante analizar estos índices para valorar el incremento que te concede la abstención. Incluso es interesante analizarlo pues puede arrebatar poder, en cuyo caso el participante debería escoger una votación en la que no se permita la abstención.

Por último, se ha concluido algo curioso. Actualmente se habla de pluralidad electoral, pero en realidad, si se analizan los datos de las elecciones se comprueba que en años pasados, por ejemplo en 1979 donde el número de partidos políticos fueron de 14, había más partidos en el Congreso de los Diputados que actualmente que hay 9. Ahora bien, estos partidos tenían un poder insignificante. Por tanto, se debería concretar más y decir que actualmente hay mayor pluralidad efectiva.

5.2. Trabajo futuro

Se plantean varias acciones a llevar a cabo para mejorar las herramientas y los cálculos.

Primeramente se plantea hacer un estudio de otros medios de comunicación para obtener información y estimar las probabilidades de coalición, dado que la teoría de *followers* de Twitter no es del todo válida por los problemas explicados a lo largo de este trabajo. Quizá una alternativa fuese incluir los Tweets, haciendo la similitud de usuarios basado en el contenido de los mismos. En el artículo [32] se presenta la teoría que discute incluir probabilidades de coalición para calcular los índices de poder de Banzhaf y Shapley-Shubik.

Las acciones referidas a los índices de poder se plantean en cuanto a optimización de código, pues en Juegos (3,2) con un número grande de participantes el tiempo de ejecución es demasiado largo. En particular, para solucionar los problemas del cálculo del índice de poder de Shapley-Shubik se propone el artículo [33], en el que se aborda una solución para aproximar linealmente el índice de Shapley-Shubik.

Por último se plantea generar más gráficas para obtener más información. Se plantean las siguientes:

- Gráficas de incremento de poder, en las que se vea como cada partido ha incrementado o disminuido su poder en las elecciones con el paso de los años.
- Gráficas de diferencias de poder entre el participante i y el participante $i+1$, pues en este trabajo esta diferencia se ha calculado a mano. Y gráficas de diferencias de poder entre juegos (3,2) configurados con mayoría simple y juegos (2,2) con mayoría absoluta para valorar que configuración le viene mejor a cada participante de manera más cómoda.

Glosario de acrónimos

- **SVP**: Sistema de votación ponderado
- **IB**: Índice de Banzhaf
- **ISS**: Índice de Shapely-Shubik
- **IBR**: Índice de Borda
- **MR**: Mariano Rajoy
- **PS**: Pedro Sánchez
- **PI**: Pablo Iglesias
- **AR**: Albert Rivera
- **PP**: Partido Popular
- **PSOE**: Partido Socialista Obrero Español
- **P**: Podemos
- **CS**: Ciudadanos
- **CIS**: Centro de Investigaciones Sociológicas

Bibliografía

[1] CIS. Encuesta preelectoral del cis. http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3140_3159/3141/es3141mar.pdf.

[2] Ministerio del Interior. Página de datos del ministerio del interior. http://www.infoelectoral.mir.es/infoelectoral/min/.

[3] Traducción de J. L. Doran y E. Hernández COMAP. *Las matemáticas en la vida cotidiana.* Addison Wesley/Universidad Autónoma de Madrid, 3 edition, 1999. Capítulo 12.

[4] Vincent Conitzer. Should social network structure be taken into account in elections? *Mathematical Social Sciences*, 64(1):100–102, 2012.

[5] Página web de la empresa dattio. http://dattio.socialnoise.com/.

[6] Página web de drago solutions. http://www.dragosolutions.com/.

[7] Documentación api twitter. https://dev.twitter.com/rest/reference/get/followers/ids. Last accessed: 30-10-2016.

[8] *Twitter vs. las urnas: así sería el Congreso de los Diputados que reflejan las redes sociales.* http://www.expansion.com/economia/politica/elecciones-generales/2016/06/14/575e8bd546163f2e2a8b463e.html. Last accessed: 2017-06-15.

[9] Página de wikipedia del conteo borda. https://es.wikipedia.org/wiki/Recuento_Borda.

[10] Ana Lucía Blas, Javier Brolo, Lorena Escobar, and Carlos González. Actualidad política. *Revista de análisis político de Guatemala*, Mayo 2014. ASIES.

[11] Página principal de asies. ://www.asies.org.gt/.

[12] Artículo sobre el sistema electoral francés. https://compolitic.wordpress.com/2010/05/19/el-sistema-electoral-frances/.

[13] Página principal del cis. http://www.cis.es/cis/opencms/ES/index.html.

[14] John F Banzhaf III. Weighted voting doesn't work: A mathematical analysis. *Rutgers L. Rev.*, 19:317, 1964.

[15] Página de wikipedia del índice de banzhaf. https://es.wikipedia.org/wiki/Indice_de_poder_de_Banzhaf.

[16] Sergio Monsalve. John nash y la teorıa de juegos. *Lecturas matemáticas*, 24:143–145, 2003.

[17] Francesc de Carreras Serra, Antonio Magaña, and Rafel Amer. Juegos simples e índice de poder de shapley-shubik. *Revista de estudios políticos*, (121):107–136, 2003.

[18] Página de wikipedia de podemos. https://es.wikipedia.org/wiki/Unidos_Podemos. Apartado: Integrantes.

[19] Edinson Montoro, Willy Barahona, Luis Macha, Gabriel Rodríguez, Emilio Castillo, Rocío De La Cruz, and Pedro Becerra. Fórmula de stirling. *Pesquimat*, 19(1), 2016.

[20] Página de wikipedia del índice de banzhaf. http://www.luschny.de/math/factorial/FastFactorialFunctions.htm.

[21] Página de wikipedia del índice de banzhaf. http://es.numberempire.com/factorialcalculator.php.

[22] Dennis Leech. Computing power indices for large voting games. *Management Science*, 49(6):831–837, 2003.

[23] John P. Lambert. *Voting Games, Power Indices, and Presidential Elections (UMAP)*. COMAP, 3 edition, 1989.

[24] Josep Freixas. Probabilistic power indices for voting rules with abstention. *Mathematical Social Sciences*, 64(1):89–99, 2012.

[25] Introducción a r. https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf.

[26] Página principal de rstudio. https://www.rstudio.com/.

[27] Página principal de shiny. https://rstudio.github.io/shinydashboard/.

[28] R para principiantes. https://cran.r-project.org/doc/contrib/rdebuts_es.pdf.

[29] Manual de r. http://www.ugr.es/~javierrp/master_files/Tutorial%20de%20R.pdf.

[30] Manual de la librería gdata. https://cran.r-project.org/web/packages/gdata/gdata.pdf.

[31] Librerías de python para trabajar con twitter. http://www.tweepy.org/.

[32] Sanoussi Bilal, Paul Albuquerque, and Madeleine O Hosli. The probability of coalition formation: spatial voting power indices. 2001.

[33] Shaheen S Fatima, Michael Wooldridge, and Nicholas R Jennings. A linear approximation method for the shapley value. *Artificial Intelligence*, 172(14):1673–1699, 2008.